Nonlinear Preconditioning: How to use a Nonlinear Schwarz Method to Precondition Newton's method

Victorita Dolean

joint work with W. Kheriji, M. J. Gander, F. Kwok, R. Masson

Dept of Mathematics and Statistics, University of Strathclyde, Glasgow

August 2, 2017

Outline

Classical vs. new approaches for non linear problems

Ideas for solving a non-linear problem $F(u) = 0$: use domain decomposition to solve the Jacobian equation in a Newton's method $\Rightarrow$ Newton-Krylov-Schwarz methods (Cai 1994, 1998).

Alternatively we can

- use the fixed point iteration $u^{n+1} = \mathcal{G}(u^n)$; to accelerate convergence, we can solve instead $\mathcal{F}(u) := \mathcal{G}(u) - u = 0$ with Newton's method.
- use a non-linear preconditioner called ASPIN introduced by Cai and Keyes (2001, 2002).

A simple example

A one dimensional non-linear model problem

$$\begin{aligned} \mathcal{L}(u) &= f, \quad \text{in } \Omega := (0, L), \\ u(0) &= 0, \\ u(L) &= 0, \end{aligned} \tag{1}$$

where $\mathcal{L}(u) = -\partial_x((1+u^2)\partial_x u)$. Two overlapping subdomains $\Omega_1 := (0,\beta)$ and $\Omega_2 := (\alpha, L)$, $\alpha < \beta$

$$\begin{aligned} \mathcal{L}(u_1^n) &= f, && \text{in } \Omega_1 := (0,\beta), \\ u_1^n(0) &= 0, \\ u_1^n(\beta) &= u_2^{n-1}(\beta), \\ \mathcal{L}(u_2^n) &= f, && \text{in } \Omega_2 := (\alpha, L), \\ u_2^n(\alpha) &= u_1^{n-1}(\alpha), \\ u_2^n(L) &= 0. \end{aligned} \tag{2}$$

Main idea for a simple problem

Global approximate solution, by glueing the approximate solutions together.

$$u^n(x) := \begin{cases} u_1^n(x) & \text{if } 0 \leq x < \frac{\alpha+\beta}{2}, \\ u_2^n(x) & \text{if } \frac{\alpha+\beta}{2} \leq x \leq L, \end{cases} \tag{3}$$

which induces two extension operators $\widetilde{P}_i$

$$u^n = \widetilde{P}_1 u_1^n + \widetilde{P}_2 u_2^n$$

Acceleration by a Newton method

Let the local solutions in the subdomains be

$$u_1^n = G_1(u^{n-1}), \qquad u_2^n = G_2(u^{n-1}), \tag{4}$$

$\Rightarrow$ the classical parallel Schwarz method can be written in compact form

$$\boxed{u^n = \sum_{i=1}^{I} \widetilde{P}_i G_i(u^{n-1}) =: \mathcal{G}_1(u^{n-1})} \tag{5}$$

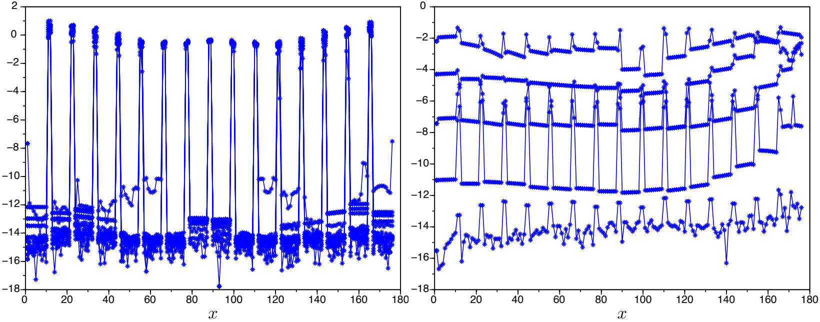
This fixed point iteration can be used as a preconditioner for Newton's method, which means to apply Newton's method to the non-linear equation

$$\tilde{\mathcal{F}}_1(u) := \mathcal{G}_1(u) - u = \sum_{i=1}^{I} \widetilde{P}_i G_i(u) - u = 0, \tag{6}$$

We call this method one level RASPEN (Restricted Additive Schwarz Preconditioned Exact Newton).

Simple tests

Forchheimer equation with 8 subdomains.



FigRAS used as a nonlinear solver, or as a preconditioner for Newton's method

The nonlinear RAS method decreases the residual only slowly at interfaces but makes it zero within the subdomains.

Comparison of ASPIN and RASPEN

Consider $F : V \to V$, V - Hilbert vector space, and the non-linear problem

$$\text{find } u \in V \text{ such that } F(u) = 0. \tag{7}$$

Let V_i be Hilbert vector spaces. Let the linear restriction and prolongation operators

$$R_i : V \to V_i, \qquad P_i : V_i \to V, \quad i = 1, \ldots, I$$

as well as the "restricted" prolongation $\widetilde{P}_i$

$$\widetilde{P}_i : V_i \to V.$$

Assumption Assume R_i and P_i satisfy for $i = 1, \ldots, I$

$$R_i P_i = I_{V_i}, \quad \text{the identity on } V_i$$

and that R_i and $\widetilde{P}_i$ satisfy

$$\boxed{\sum_{i=1}^{I} \widetilde{P}_i R_i = I_V}$$

Formulation of ASPIN and RASPEN

Define the local inverse $G_i : V \to V_i$ to be solutions of

$$R_i F\left(P_i G_i(u) + (I - P_i R_i) u\right) = 0, \tag{8}$$

Then, one level RASPEN solves the non-linear equation

$$\boxed{\tilde{\mathcal{F}}_1(u) = \sum_{i=1}^{I} \widetilde{P}_i G_i(u) - u} \tag{9}$$

using Newton's method. The preconditioned nonlinear function (9) corresponds to the fixed point iteration

$$\boxed{u^n = \sum_{i=1}^{I} \widetilde{P}_i G_i(u^{n-1})} \tag{10}$$

Formulation of ASPIN and RASPEN

For $u \in V$ define the corrections $C_i(u) \in V_i$ such that

$$R_i F(u + P_i C_i(u)) = 0, \quad i = 1, \cdots, I. \tag{11}$$

where

$$G_i(u) = R_i u + C_i(u).$$

Then, the one level ASPIN is

$$\boxed{\mathcal{F}_1(u) = \sum_{i=1}^{I} P_i C_i(u) = \sum_{i=1}^{I} P_i G_i(u) - \sum_{i=1}^{I} P_i R_i u} \tag{12}$$

This corresponds to the non-linear fixed point iteration

$$\boxed{u^n = u^{n-1} - \sum_{i=1}^{I} P_i R_i u^{n-1} + \sum_{i=1}^{I} P_i G_i(u^{n-1})} \tag{13}$$

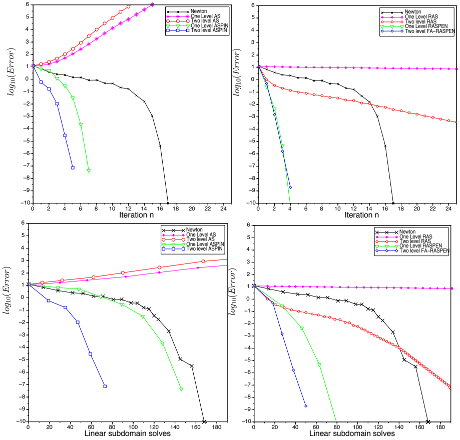
ASPIN vs. RASPEN

Remarks:

- The iterative version of ASPIN is not convergent in the overlap, and needs a relaxation parameter to yield convergence for the non-linear case.
- The use of restricted extension makes the RASPEN iterative method convergent.
- The only interest in the additive correction in the overlap is that in the linear case for a symmetric problem, the preconditioner remains symmetric.

Numerical comparison

Forchheimer equation on a domain of unit size, 8 subdomains, $h = 1/100$.



Jacobian matrices ASPIN/RASPEN

Denote by

$$u^{(i)} := P_i G_i(u) + (I - P_i R_i)u \quad \text{and} \quad J(v) := \frac{dF}{du}(v) \tag{14}$$

By differentiating the subdomain solves we get

$$\frac{dG_i}{du}(u) = -(R_i J(u^{(i)}) P_i)^{-1} R_i J(u^{(i)}) + R_i.$$

Jacobian of RASPEN

$$\boxed{\frac{d\tilde{\mathcal{F}}_1}{du}(u) = \sum_{i=1}^{I} \widetilde{P}_i \frac{dG_i}{du}(u) - I = -\sum_{i=1}^{I} \widetilde{P}_i (R_i J(u^{(i)}) P_i)^{-1} R_i J(u^{(i)})} \tag{15}$$

Similarly, for ASPEN we get

$$\boxed{\frac{d\mathcal{F}_1}{du}(u) = \sum_{i=1}^{I} P_i \frac{dG_i}{du}(u) - \sum_{i=1}^{I} P_i R_i = -\sum_{i=1}^{I} P_i (R_i J(u^{(i)}) P_i)^{-1} R_i J(u^{(i)})} \tag{16}$$

In ASPIN, this exact Jacobian is replaced by the inexact Jacobian

$$\frac{d\mathcal{F}_1}{du}^{inexact}(u) = -\left(\sum_{i=1}^{I} P_i (R_i J(u) P_i)^{-1} R_i\right) J(u).$$

Remarks:

- The inexact Jacobian corresponds to the Jacobian $J(u)$ of $F(u)$ preconditioned by the restricted additive Schwarz preconditioner, up to the minus sign.
- The exact Jacobian is however also easily accessible, since the non linear Newton solver for the non linear subdomain system $R_i F\left(P_i G_i(u) + (I - P_i R_i)u\right) = 0$ already computes the Jacobian matrix $R_i J(u^{(i)}) P_i$.
- It suffices to compute instead the matrix $R_i J(u^{(i)})$ for each non linear subdomain system to easily obtain the exact Jacobian of $\mathcal{F}_1$.

Coarse spaces for RASPEN and ASPIN

Let V_0 and the linear restriction and prolongation operators

$$R_0 : V \to V_0 \quad \text{and} \quad P_0 : V_0 \to V. \tag{17}$$

We introduce a projection operator in the residual space

$$\widetilde{R}_0 : V' \to V_0'. \tag{18}$$

Let $F_0 : V_0 \to V_0$ be the coarse non-linear function,

$$F_0(u_0) = \widetilde{R}_0 F(P_0(u_0)). \tag{19}$$

Let the coarse solution for two level ASPIN be $u_0^* \in V_0$, i.e. $F_0(u_0^*) = 0$.

Coarse correction in two-level ASPIN

The coarse correction $C_0^A : V \to V_0$ is defined by

$$F_0(C_0^A(u) + u_0^*) = -\widetilde{R}_0 F(u), \tag{20}$$

and the associated two level ASPIN function uses the coarse correction in an additive fashion, i.e. Newton's method is used to solve

$$\mathcal{F}_2(u) = P_0 C_0^A(u) + \sum_{i=1}^{I} P_i C_i(u) = 0. \tag{21}$$

This corresponds to the non-linear two level fixed point iteration

$$\boxed{u^{n+1} = u^n + P_0 C_0^A(u^n) + \sum_{i=1}^{I} P_i C_i(u^n)}$$

which is not convergent without relaxation parameter and also slows down the Newton solver.

Coarse correction in two-level FAS-RASPEN

Use the well established non-linear coarse correction $C_0(u)$ from the full approximation scheme

$$F_0(C_0(u) + R_0 u) = F_0(R_0 u) - \widetilde{R}_0 F(u). \tag{22}$$

This gives a different coarse correction from ASPIN and this coarse correction is used in a multiplicative fashion in RASPEN, i.e. we solve with Newton the preconditioned non-linear system

$$\widetilde{\mathcal{F}}_2(u) = P_0 C_0(u) + \sum_{i=1}^{n} \widetilde{P}_i C_i(u + P_0 C_0(u)) = 0. \tag{23}$$

This corresponds to the non-linear two level fixed point iteration

$$\boxed{u^{n+1} = u^n + P_0 C_0(u^n) + \sum_{i=1}^{n} \widetilde{P}_i C_i(u^n + P_0 C_0(u^n))}$$

This iteration is convergent.

Forchheimer model

Let the Forchheimer parameter $\beta > 0$, the permeability $\lambda \in L^\infty(\Omega)$ such that $0 < \lambda_{min} \leq \lambda(x) \leq \lambda_{max}$ for all $x \in \Omega$, and the function $q(g) = \text{sgn}(g)\frac{-1+\sqrt{1+4\beta|g|}}{2\beta}$. The Forchheimer model on the interval $\Omega = (0, L)$ is defined by the equation

$$\left\{\begin{array}{rcll} (q(-\lambda(x)u(x)'))' & = & f(x) & \text{in } \Omega, \\ u(0) & = & u_0^D, & \\ u(L) & = & u_L^D. & \end{array}\right. \tag{24}$$

Note that at the limit when $\beta \to 0^+$, we recover the linear Darcy equation.

Notations

Each Newton iteration requires two major steps:

1. Evaluation of the fixed point function $\mathcal{F} \to$ solving a non-linear problem in each subdomain $\to$ maximum number of inner iterations needed by the subdomains at the outer iteration j: ls_j^{in}.
2. The Jacobian matrix needs to be inverted (GMRES) $\to$ a linear subdomain solve per subdomain per GMRES iteration $\to$ the number of linear solves needed by GMRES at the outer Newton iteration step j: ls_j^G.

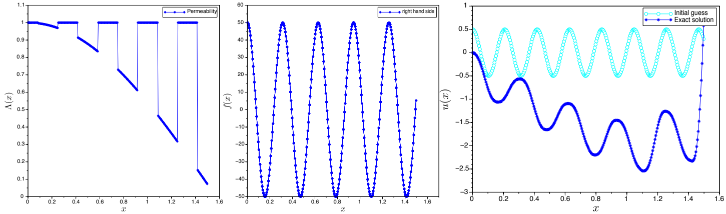
The total number of linear subdomain solves after n outer Newton iterations

$$\boxed{LS_n := \sum_{j=1}^{n} \left(ls_j^{in} + ls_j^G \right)}$$

- PIN $\to$ outer Newton iterations for ASPIN
- PEN $\to$ outer Newton iterations for RASPEN.

Test case: a smooth vs a non smooth example

- $\Omega = (0,1)$, boundary conditions $u(0) = 0$ and $u(1) = 1$, $\beta = 1$.
- Linear stopping criterion for GMRES: 10^{-8}, non linear stopping criteria for the inner and outer Newton iterations: 10^{-8}.
- Smooth example: $\lambda(x) = \cos x$, $f(x) = \cos x$.



FigNon smooth example: $\lambda(x)$ (left), $f(x)$ (middle), initial guess/solution (right).

ASPIN						
No of domains	10		20		40	
Iter type / Overlap	PIN	LS_n	PIN	LS_n	PIN	LS_n
h	5	118	5	228	6	520
3h	5	118	5	227	6	516
5h	5	117	5	222	6	480

RASPEN						
No of domains	10		20		40	
Iter type / Overlap	PEN	LS_n	PEN	LS_n	PEN	LS_n
h	4	92	4	172	4	340
3h	4	87	4	172	4	331
5h	4	88	4	168	4	313

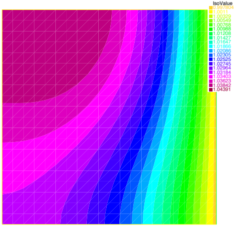
Two level ASPIN						
No of domains	10		20		40	
Iter type / Overlap	PIN	LS_n	PIN	LS_n	PIN	LS_n
h	5	140	5	240	5	280
3h	5	130	6	170	6	200
5h	5	115	7	149	6	147

Two level FAS RASPEN						
No of domains	10		20		40	
Iter type / Overlap	PEN	LS_n	PEN	LS_n	PEN	LS_n
h	4	77	3	87	4	131
3h	3	60	3	67	4	90
5h	3	55	3	57	3	57

ASPIN						
No. of domains	10		20		40	
Overlap \ Iter type	PIN	LS_n	PIN	LS_n	PIN	LS_n
h	8	184	15	663	-	-
3h	7	156	14	631	11	883
5h	6	130	11	479	10	744
RASPEN						
No of domains	10		20		40	
Overlap \ Iter type	PEN	LS_n	PEN	LS_n	PEN	LS_n
h	7	150	9	369	9	701
3h	7	145	8	324	9	691
5h	6	126	7	274	9	659
Two-level ASPIN						
No of domains	10		20		40	
Overlap \ Iter type	PIN	LS_n	PIN	LS_n	PIN	LS_n
h	7	184	9	316	8	285
3h	6	141	9	246	7	183
5h	6	135	8	199	7	164
Two-level FAS-RASPEN						
No of domains	10		20		40	
Overlap \ Iter type	PEN	LS_n	PEN	LS_n	PEN	LS_n
h	7	134	9	272	8	258
3h	7	133	8	220	6	136
5h	6	112	8	211	6	116

Two-dimensional examples

A non-linear diffusion problem discretised by P1 finite elements on a uniform triangular mesh (computations using FreeFEM++)



$$\left\{ \begin{array}{rcl} -\nabla \cdot ((1+u^2)\nabla u) & = & f, \quad \Omega = [0,1]^2, \\ u & = & 1, \quad x = 1, \\ \dfrac{\partial u}{\partial \mathbf{n}} & = & 0, \quad \text{otherwise.} \end{array} \right. \tag{25}$$

Detailed convergence RASPEN vs. ASPIN

- decomposition into $N \times N$ subdomains with an overlap of one mesh size h
- number of degrees of freedom per subdomain fixed

		1-Level				2-Level			
$N \times N$	n	ls_n^G	ls_n^{in}	$ls_n^{\min}$	LS_n	ls_n^G	ls_n^{in}	$ls_n^{\min}$	LS_n
2×2	1	15(20)	4(4)	3(3)		13(23)	4(4)	3(3)	
	2	17(23)	3(3)	3(3)	59(78)	15(26)	3(3)	3(3)	54(86)
	3	18(26)	2(2)	2(2)		17(28)	2(2)	2(2)	
4×4	1	32(37)	3(3)	3(3)		18(33)	3(3)	3(3)	
	2	35(41)	3(3)	2(2)	113(132)	22(39)	3(3)	2(2)	74(126)
	3	38(46)	2(2)	2(2)		26(46)	2(2)	2(2)	
8×8	1	62(71)	3(3)	2(2)		18(35)	3(3)	3(2)	
	2	67(77)	3(3)	2(2)	211(240)	23(44)	3(3)	2(2)	77(139)
	3	74(84)	2(2)	1(2)		28(53)	2(2)	2(1)	
16×16	1	125(141)	3(3)	2(2)		18(35)	3(3)	3(2)	
	2	136(155)	2(2)	2(2)	418(471)	23(44)	2(2)	2(2)	75(140)
	3	150(167)	2(2)	1(1)		27(54)	2(2)	2(1)	

$\Rightarrow$ RASPEN clearly outperforms ASPIN (which not convergent as a basic fixed point iteration).

Conclusions

- accelerate fixed point iterations for non-linear problems using Newton's method $\rightarrow$ guiding principle for constructing non-linear preconditioners.
- explore the parallel properties of the RASPEN method on more realistic models and configurations.
- design of other nonlinear preconditioners (ongoing work) based on non-overlapping decompositions using Neumann-Neumann or Robin-Robin iterations. (application to the Richards equation).

References and acknowledgements

Acknowledgements: TOTAL S.A. (Oil Company), Hong Kong Research Grant Council (Grant No. ECS/22300115), NSFC Young Scientist Fund (Grant No. 11501483).

Bibiography:

V. Dolean, P. Jolivet, F. Nataf, "An Introduction to Domain Decomposition Methods: algorithms, theory and parallel implementation", SIAM, 2015, http://bookstore.siam.org/ot144/.

V Dolean, MJ Gander, W Kheriji, F Kwok, R Masson, "Nonlinear preconditioning: how to use a nonlinear Schwarz method to precondition Newton's method", vol 38(6), pp. A3357-A3380, SIAM Journal on Scientific Computing, 2016.