Comparison of Projection Methods derived from Deflation, Domain Decomposition and Multigrid Methods

Reinhard Nabben
TU Berlin

Y.A. Erlangga TU Berlin, K. Vuik, TU Delft, J. Tang TU Delft

supported by

Deutsche Forschungsgemeinschaft (DFG)

20.08.2007

Outline

Deflation

Comparison Deflation vs. additive coarse grid corrections

Deflation vs balancing

Comparison of Projection methods

Numerical comparison

Conclusion

$$Ax = b$$

A is sparse and symmetric positive definite; condition number of A is huge, due to large jumps in the coefficients

Applications:

- reservoir simulations
- porous media flow
- elasticity
- more

Deflated CG

Nicolaides 1987, Mansfield 1988, 1990, Kolotilina 1998, Vuik, Segal, and Meijerink 1999, Morgan 1995, Saad, Yeung, Erhel, and Guyomarch 2000, Frank and Vuik 2001, Blaheta 2006

Deflation and restarted GMRES

Morgan 1995, Erhel, Burrage, and Pohl 1996, Chapman and Saad 1997, Eiermann, Ernst, and Schneider 2000, Morgan 2002

Clemens et al. 2003,2004, de Sturler et al. 2006, Aksoylu, H. Klie, and M.F. Wheeler 2007

Deflation with eigenvectors

$$Au_i = \lambda_i u_i \qquad Z = [u_1, \ldots, u_r] \qquad u_i^T u_j = \delta_{ij}$$

$$P = I - AZ(Z^T AZ)^{-1} Z^T, \quad Z \in \mathbb{R}^{n \times r},$$

$$spectrum(PA) = \{0, \ldots, 0, \lambda_{r+1}, \ldots \lambda_n\}$$

Deflation with general vectors

$$Z = [z_1, \ldots, z_r] \qquad rankZ = r \quad E = Z^T A Z$$

$$P = I - AZE^{-1}Z^T, \quad Z \in \mathbb{R}^{n \times r},$$

$$PAZ = 0$$

$$spectrum(PA) = \{0, \ldots, 0, \mu_{r+1}, \ldots \mu_n\}$$

Deflation for linear systems

$$Z \in \mathbb{R}^{n \times r} \quad Z = [z_1, \ldots, z_r] \quad rank Z = r$$

$$Ax = b \qquad P = I - AZE^{-1}Z^T$$

We have: $x = (I - P^T)x + P^T x$ Compute both!

1. $(I - P^T)x = Z(Z^T AZ)^{-1} Z^T b$
2. Solve $PA\tilde{x} = Pb$ preconditioner M^{-1}:
   $M^{-1}PA\tilde{x} = M^{-1}Pb$
3. Build $P^T\tilde{x} = P^T x$

Another Deflation Variant (Kolotilina; Saad et al.)

Choose a random vector, $\bar{x}$,

start CG with $x_0 := Z(Z^T AZ)^{-1} Z^T b + P^T \bar{x}$

for the system

$$P^T M^{-1} Ax = P^T M^{-1} b,$$

Deflation

M^{-1} preconditioner, ILU $\quad$ Z appro. eigenvectors

$ZE^{-1}Z^T$

Domain decomposition

M^{-1} add. Schwarz $\quad$ Z grid transfer operator

$ZE^{-1}Z^T$ coarse grid correction

Multigrid

M^{-1} smoother $\quad$ Z grid transfer operator

$ZE^{-1}Z^T$ coarse grid correction

Additive coarse grid corrections

$Z^T : \mathbb{R}^n \to \mathbb{R}^r :$ restriction $\quad Z : \mathbb{R}^r \to \mathbb{R}^n :$ prolongation

$Z^T AZ$ Galerkin product

$Z(Z^T AZ)^{-1} Z^T$ Coarse Grid Correction

Preconditioner

$$P_{ad} = M^{-1} + Z(Z^T AZ)^{-1} Z^T$$

Bramble, Pasciak and Schatz 1986, Dryja and Widlund 1990

Comparison: Deflation vs. additive coarse grid correction

$$
\begin{aligned}
M^{-1}P &= M^{-1} - M^{-1}AZE^{-1}Z^T \\
P_{ad} &= M^{-1} + ZE^{-1}Z^T
\end{aligned}
$$

Theorem

Nabben, Vuik 04
For all Z with $rank Z = r$ we have:

$$
\begin{aligned}
\lambda_n(M^{-1}PA) &\leq \lambda_n(P_{ad}A) \\
\lambda_{r+1}(M^{-1}PA) &\geq \lambda_1(P_{ad}A)
\end{aligned}
$$

Thus:

$$cond_{eff}(M^{-1}PA) \leq cond(P_{ad}A)$$

abstract Balancing Neumann-Neumann preconditioner

Mandel 1993, Dryja and Widlund 1995, Mandel and Brezina 1996

M^{-1} Neumann-Neumann preconditioner

$$P_B = (I - ZE^{-1}Z^TA)M^{-1}(I - AZE^{-1}Z^T) + ZE^{-1}Z^T,$$

$$E = Z^TAZ, \qquad Z \in \mathbb{R}^{n \times r}$$

$$P = I - AZE^{-1}Z^T,$$

$$P_B = P^TM^{-1}P + ZE^{-1}Z^T.$$

Deflation and abstract balancing preconditioner

Theorem

Nabben, Vuik 06

- $cond_{eff}(M^{-1}PA) \leq cond(P_B A)$

Deflation and abstract balancing preconditioner

Theorem

Nabben, Vuik 06

- $cond_{eff}(M^{-1}PA) \leq cond(P_B A)$
- $$spectrum(M^{-1}PA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}$$
  $$spectrum(P_B A) = \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\}$$

Deflation and abstract balancing preconditioner

Theorem

Nabben, Vuik 06

- $cond_{eff}(M^{-1}PA) \leq cond(P_BA)$
- $$\begin{aligned} spectrum(M^{-1}PA) = & \quad \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\} \\ spectrum(P_BA) = & \quad \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\} \end{aligned}$$
- For $\tilde{x}_{0,D} = x_{0,B}$ $\qquad \|x - x_{k,D}\|_A \leq \|x - x_{k,B}\|_A.$

Deflation and abstract balancing preconditioner

Theorem

Nabben, Vuik 06

- $cond_{eff}(M^{-1}PA) \leq cond(P_B A)$
- $$spectrum(M^{-1}PA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}$$
  $$spectrum(P_B A) = \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\}$$
- For $\tilde{x}_{0,D} = x_{0,B}$ $\quad \|x - x_{k,D}\|_A \leq \|x - x_{k,B}\|_A$.
- $\tilde{x}_{0,D} = \bar{x}$ and $x_{0,B} = ZE^{-1}Z^T b + P^T \bar{x}$ then

$$x_{k,D} = x_{k,B}.$$

Implementation of the Balancing preconditioner

Solve $Ax = b$
Take $x_{0,B} = ZE^{-1}Z^Tb + P^T\bar{x}$

Then the balancing preconditioner $P^TM^{-1}P + ZE^{-1}Z^T$ can be implemented with the use of

$$P^TM^{-1}$$

only.

Mandel 93, Toselli, Widlund 04.

Motivation: Save of work per iteration

Our more detailed analysis shows

- better effective condition number
- better clustering
- better A-norm of the error

Remarks from others

- Balancing is robust w.r.t. inexact solves, deflation not.

Comparison of Projection methods

Deflation, Variant 1

$$M^{-1}P$$

Deflation, Variant 2

$$P^T M^{-1}$$

Balancing method

$$P^T M^{-1} P + Z E^{-1} Z^T$$

Reduced balancing method, Variant 1

$$P^T M^{-1} P$$

Reduced balancing method, Variant 2

$$P^T M^{-1}$$

Adapted Deflation, Variant 1

$$M^{-1}P + ZE^{-1}Z^T$$

Multigrid: results from non-symmetric multigrid
first coarse grid correction, then smoothing

Adapted Deflation, Variant 2

$$P^TM^{-1} + ZE^{-1}Z^T$$

Multigrid: results from non-symmetric multigrid
first smoothing, then coarse grid correction

Name	**Method**	**Operator**
PREC	Traditional Preconditioned CG	M^{-1}
AD	Additive Coarse Grid Correc.	$M^{-1} + Q$
DEF1	Deflation Variant 1	$M^{-1}P$
DEF2	Deflation Variant 2	$P^T M^{-1}$
A-DEF1	Adapted Deflation Variant 1	$M^{-1}P + Q$
A-DEF2	Adapted Deflation Variant 2	$P^T M^{-1} + Q$
BNN	Abstract Balancing	$P^T M^{-1} P + Q$
R-BNN1	Reduced Balancing Variant 1	$P^T M^{-1} P$
R-BNN2	Reduced Balancing Variant 2	$P^T M^{-1}$

$$Q = ZE^{-1}Z^T = Z(Z^T AZ)^{-1}Z^T$$

Method	$\mathcal{V}_{\text{start}}$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{V}_{\text{end}}$
PREC	$\bar{x}$	M^{-1}	I	I	x_{j+1}
AD	$\bar{x}$	$M^{-1} + Q$	I	I	x_{j+1}
DEF1	$\bar{x}$	M^{-1}	I	P	$Qb + P^T x_{j+1}$
DEF2	$Qb + P^T \bar{x}$	M^{-1}	P^T	I	x_{j+1}
A-DEF1	$\bar{x}$	$M^{-1}P + Q$	I	I	x_{j+1}
A-DEF2	$Qb + P^T \bar{x}$	$P^T M^{-1} + Q$	I	I	x_{j+1}
BNN	$\bar{x}$	$P^T M^{-1} P + Q$	I	I	x_{j+1}
R-BNN1	$Qb + P^T \bar{x}$	$P^T M^{-1} P$	I	I	x_{j+1}
R-BNN2	$Qb + P^T \bar{x}$	$P^T M^{-1}$	I	I	x_{j+1}

- Select random $\bar{x}$ and $\mathcal{V}_{\text{start}}, \mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{V}_{\text{end}}$ from Table
- $x_0 := \mathcal{V}_{\text{start}}, \;\; r_0 := b - Ax_0$
- $z_0 := \mathcal{M}_1 r_0, \; p_0 := \mathcal{M}_2 z_0$
- FOR $j := 0, 1, \ldots,$ until convergence
  - $w_j := \mathcal{M}_3 A p_j$
  - $\alpha_j := (r_j, z_j)/(p_j, w_j)$
  - $x_{j+1} := x_j + \alpha_j p_j$
  - $r_{j+1} := r_j - \alpha_j w_j$
  - $z_{j+1} := \mathcal{M}_1 r_{j+1}$
  - $\beta_j := (r_{j+1}, z_{j+1})/(r_j, z_j)$
  - $p_{j+1} := \mathcal{M}_2 z_{j+1} + \beta_j p_j$
- ENDFOR
- $x_{\text{it}} := \mathcal{V}_{\text{end}}$

Theorem

Tang, Nabben, Vuik, Erlangga 07
The spectrum of the systems preconditioned by DEF1, DEF2, R-BNN1 or R-BNN2 is given by

$$\{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}.$$

The spectrum of the systems preconditioned by A-DEF1, A-DEF2, BNN is given by

$$\{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\}.$$

Theorem

Tang, Nabben, Vuik, Erlangga 07

With starting vector $x_0 = Qb + P^T\bar{x}$

BNN, DEF2, A-DEF2, R-BNN1 and R-BNN2, are identical in exact arithmetic.

Lemma

Suppose that $x_0 = Qb + P^T\bar{x}$ *is used as starting vector.*

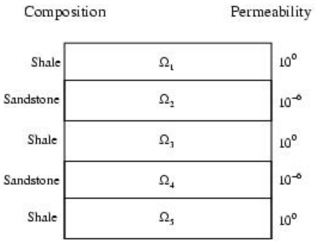
- $Qr_{j+1} = \mathbf{0}$;
- $Pr_{j+1} = r_{j+1}$,

for all $j = 0, 1, 2, \ldots$.

We consider the Poisson equation with a discontinuous coefficient,

$$-\nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}) \right) = \mathbf{0}, \quad \mathbf{x} = (x, y) \in \Omega = (0,1)^2, \quad (1)$$

where ρ and p denote the piecewise-constant density and fluid pressure, respectively. The contrast, $\epsilon = 10^{-6}$, is fixed, which is the jump between the high and low density.

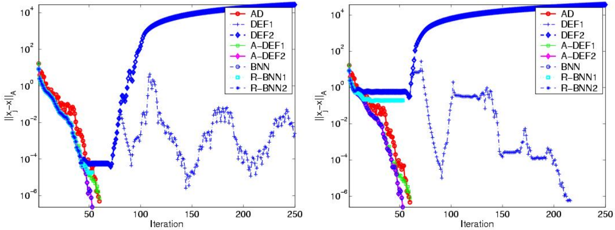


Geometry of the porous media problem with $r = 5$ layers having a fixed density ρ. The number of deflation vectors and layers is equal.

Experiment using Inaccurate Coarse Solves

$$\widetilde{E}^{-1} := (I + \psi R)E^{-1}(I + \psi R), \quad \psi > 0, \qquad (2)$$

where $R \in \mathbb{R}^{r \times r}$ is a symmetric random matrix with elements from the interval $[-0.5, 0.5]$



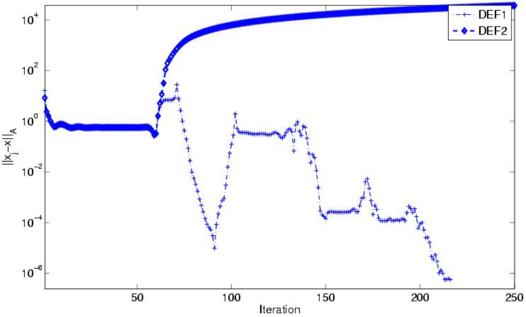
$\psi = 10^{-12}$ $\psi = 10^{-8}$

Exact errors in the $A-$norm for the test problem with $n = 29^2, r = 5$ and $\widetilde{E}^{-1}$.

Experiment using Inaccurate Coarse Solves

Comparison of DEF1 and DEF2

DEF1 : $M^{-1}P$ DEF2 : $P^T M^{-1}$ (Saad et al.)

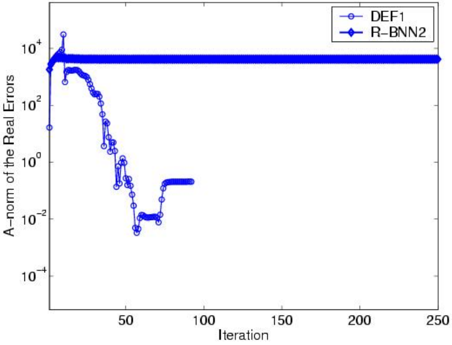


$\psi = 10^{-8}$

Experiment using Inaccurate Coarse Solves

Comparison of DEF1 and R-BNN2

DEF1 : $M^{-1}P$ R-BNN2 : P^TM^{-1} (Widlund et al.)

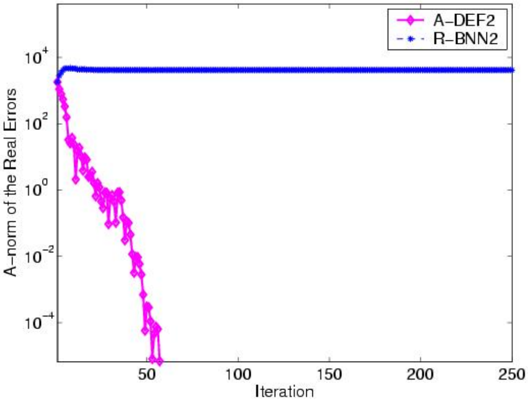


$\psi = 10^{-4}$

Experiment using Inaccurate Coarse Solves

Comparison of A-DEF2 and R-BNN2

A-DEF2 : $P^T M^{-1} + Q$ R-BNN2 : $P^T M^{-1}$



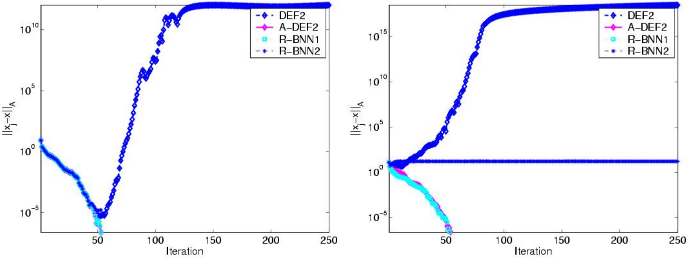
$\psi = 10^{-8}$

Experiment using Perturbed Starting Vectors

Reinhard Nabben
TU Berlin

$$\mathcal{W}_{\text{start}} := \gamma y_0 + \mathcal{V}_{\text{start}}, \quad \gamma \geq 0,$$

where y_0 is a random vector with elements from the interval $[-0.5, 0.5]$.



$\gamma = 10^{-6}$ $\gamma = 10^{0}$

Exact errors in the $A-$norm for the test problem with $n = 29^2, r = 5^2$ and perturbed starting vectors.

Conclusion of the experiments

$$P^T M^{-1} + Q \quad \text{with starting vector } \; x_0 = Qb + P^T \bar{x}$$

is robust w.r.t. inexact coarse grid solves
is robust w.r.t. perturbed starting vector

Conclusion

- We gave a comparison of projection methods derived from deflation, domain decomposition and multigrid methods
- We proved a number of theoretical comparisons and performed a number of experiments
- If one want to choose between deflation variants, DEF1 seems to be better
- Optimal implementation of BNN is as sensitive as deflation w.r.t. inexact solves
- $P^T M^{-1} + Q$ is a robust deflation variant

More: Non-symmetric Problems

- Erlangga, Nabben 06:

$$Z^T \to Y^T \qquad E \to Y^T AZ$$

$$P_D = I - AZE^{-1}Y^T$$

$$P_D^T \to Q_D = I - ZE^{-1}Y^T A$$

$$P_B = Q_D M^{-1} P_D + ZE^{-1}Y^T$$

$$\|M^{-1}(b - Au_{k,D})\|_2 \leq \|M^{-1}(b - Au_{k,B})\|_2.$$

Non-symmetric Problems

- Erlangga, Nabben 07a:

  Multilevel Deflation; Multilevel Projection Krylov method

  $$P_N = P_D + \lambda Z E^{-1} Y^T$$

  outer and inner iterations: FGMRES
- Erlangga, Nabben 07b:
  Multilevel Projection Krylov Method for the Helmholtz equation

http://www.math.tu-berlin.de/˜nabben