Aalto University
School of Science and Technology

Approximate factoring of the inverse

Mikko Byckling, Marko Huhtanen

Introduction

Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm

Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE

Conclusions

Approximate factoring of the inverse

Mikko Byckling   Marko Huhtanen

Aalto University School of Science and Technology, Finland
mikko.byckling@tkk.fi

Sparse Days, 15.6.2010

Contents

1 Introduction

2 Computing approximate factors for the inverse

3 Numerical experiments

4 Conclusions

Introduction I

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

Let $A \in \mathbb{C}^{n \times n}$ be large, sparse and nonsingular matrix.

Let $\mathcal{W}$ and $\mathcal{V}_1$ be sparse matrix subspaces of $\mathbb{C}^{n \times n}$ with nonsingular elements of $\mathcal{V}_1$ being readily invertible.

To approximately factor A into product WV_1^{-1}, consider

$$AW \approx V_1, \tag{1}$$

with non-zero matrices $W \in \mathcal{W}$ and $V_1 \in \mathcal{V}_1$ regarded as variables **both**.

Introduction II

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

Define a linear operator $L : \mathcal{W} \longmapsto \mathbb{C}^{n \times n}$ as

$$W \longmapsto LW = (I - P_1)AW, \tag{2}$$

where P_1 denotes orthogonal projection on $\mathbb{C}^{n \times n}$ onto $\mathcal{V}_1$.

If W is in the nullspace of L, then $AWV_1^{-1} = I$ is a factorization of the inverse with $V_1 = P_1AW = AW$.

Introduction III

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

With these, we have $AW \approx V_1 = P_1 AW$ if and only if $(I - P_1)AW \approx 0$.

This leads to optimality criterion

$$\min_{W \in \mathcal{W}, ||W||_F = 1} ||(I - P_1)AW||_F \,, \tag{3}$$

for generating approximate factors W and V_1 in terms of the singular values of the linear map (2).

Approximate factors

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

To solve the minimization problem (3), we need approximately compute the smallest singular values of the linear operator (2).

Adjoint operator of L is

$$P_{\mathcal{W}} A^*(I - P_1) : \mathbb{C}^{n \times n} \to \mathcal{W},$$

where $P_{\mathcal{W}}$ denotes orthogonal projection on $\mathbb{C}^{n \times n}$ onto $\mathcal{W}$.

We numerically approximate the smallest eigenpairs of

$$W \longmapsto L^* L W = P_{\mathcal{W}} A^*(I - P_1) A W.$$

To this end, we apply power method to operator

$$\alpha I - P_{\mathcal{W}} A^*(I - P_1) A \text{ on } \mathcal{W},$$

where $\alpha \in \mathbb{R}^+$ is chosen as $\alpha = r||A||_2^2$, $1/2 < r \leq 3/4$.

Power method for computing approximate factors

Approximate factoring of the inverse

Mikko Byckling, Marko Huhtanen

Introduction

Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm

Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE

Conclusions

Power method for computing approximate factors

Set $\alpha \in \mathbb{R}^+$ and choose an initial factor $W \in \mathcal{W}$
repeat
    $M = (I - P_1)AW$
    $N = P_{\mathcal{W}} A^* M$
    $W := \alpha W - N$
    $W := W / ||W||_F$
until stopped

Once the iteration is stopped V_1 is acquired by computing $V_1 = P_1 AW$.

Implementation

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

Let $\mathcal{W}$ and $\mathcal{V}_1$ be standard matrix subspaces, i.e., the sparsity patterns of $\mathcal{W}$ and $\mathcal{V}_1$ determine their dimension.

Then projectors P_1 and $P_\mathcal{W}$ can be regarded as acting columnwise and the whole algorithm is **readily parallelizable**.

Denote by w_j the jth column of W and by $(I - P_1)_j$, $(I - P_\mathcal{W})_j$ and $P_{\mathcal{W},j}$ projections of the jth column.

To have $m_j = (I - P_1)_j A w_j$, only entries appearing in the *complement* of $\mathcal{V}_1$ on the jth column have to be computed.

To have $n_j = P_{\mathcal{W},j} A^* m_j$, only entries appearing on the jth column of $\mathcal{W}$ have to be computed.

Finally $\alpha w_j - n_j$ overwrites w_j

Computational complexity

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

Let $\mathcal{W}$ and $\mathcal{V}_1$ be standard matrix subspaces.

Assume $\mathrm{nz}(A) = O(k_1 n)$, $\mathrm{nz}(\mathcal{W}) = O(k_2 n)$ and $\mathrm{nz}(\mathcal{V}_1) = O(k_3 n)$ with $k_1, k_2, k_3 \ll n$.

Then one iteration round requires $O((2k_1 + 5)k_2 n)$ operations.

Choosing subspaces $\mathcal{W}$ and $\mathcal{V}_1$

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

Sparsity structures for subspaces $\mathcal{W}$ and $\mathcal{V}_1$ can be adaptively constructed with numerical dropping.

Dropping leads to sparsity structures that greatly differ from the sparsity structures of A, A^* or A^*A.

Let $v \in \mathbb{C}^n$ with the entries v_j. Numerically drop entries by

- relative tolerance τ. Store only entries v_j for which $|v_j| \geq \tau_j = \tau||v||_2$ holds.
- count p. Store only the p largest elements of v.

Usually $\tau \approx 1e-5$ and $p \approx 10$.

Numerical dropping can be computationally expensive.

Power method for computing approximate factors

Power method for computing approximate factors

Let $\alpha \in \mathbb{R}^+$ and $W = [w_1, \ldots, w_n] \in \mathcal{W}$
repeat
  for all columns $j = 1, \ldots, n$ **in parallel do**
    $m_j = (I - P_1)_j A w_j$
    Find large entries of $(I - P_{\mathcal{W}})_j A^* m_j$, update $\mathcal{W}$
    $n_j := P_{\mathcal{W},j} A^* m_j$
    $w_j := \alpha w_j - n_j$
    Apply numerical dropping to w_j and update $\mathcal{W}$
  end for
  $W := W / ||W||_F$
until stopped
Compute $V_1 = P_1 A W$ and apply numerical dropping to V_1

Numerical experiments

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

We iteratively solve the preconditioned linear system

$$AMy = b \text{ with } x = My, \tag{4}$$

with $b \in \mathbb{C}^n$ and matrix $A \in \mathbb{C}^{n \times n}$ that is large and sparse. With approximate factors, we have $M = WV_1^{-1}$.

Right-hand side is $b = \hat{b}/||\hat{b}||_2$ where $\hat{b} = (1, 1, \ldots, 1)$. Initial guess was set as $x_0 = (0, 0, \ldots, 0)$.

Iteration was considered converged when

$$\frac{||b - Ax_m||_2}{||b||_2} \leq 10^{-6}.$$

Subspace $\mathcal{V}_1$ was chosen to be a set of block diagonal matrices with blocksize k.

Matrix sherman2 I

Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

Let A to be the $\mathrm{sherman2}$ from Matrix Market arising in oil reservoir simulation having $n = 1080$, $\mathrm{nz}(A) = 23094$ and $\kappa(A) = 1.4E + 12$.

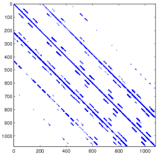
Use $\mathrm{GMRES}(30)$ to solve the preconditioned linear system (4).

Determine the sparsity pattern of $\mathcal{W}$ using numerical dropping with $\tau = 1e - 10$ and $p = 10$ and initial guess $W_0 = I$.

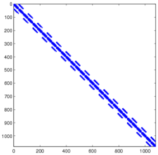
Choose $\mathcal{V}_1$ as the set of block diagonal matrices with blocksize $k = 72$. Use numerical dropping with $\tau = 1e - 10$ and $p = 15$.

Aalto University
School of Science
and Technology

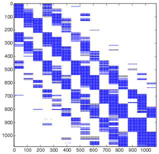
Matrix sherman2 II



(a) W



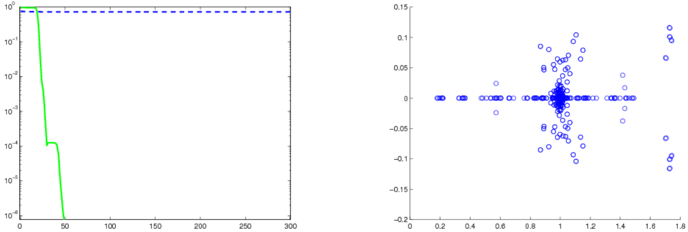
(b) V_1



(c) WV_1^{-1}

Figure: Sparsity patterns of the factors W and V_1 and the preconditioner WV_1^{-1} after 5 iterations. Number of nonzeroes $\mathrm{nz}(W)=9204$, $\mathrm{nz}(V_1)=12190$ and $\mathrm{nz}(WV_1^{-1})=198837$.

Matrix sherman2 III



(a) (b)

Figure: (a) Convergence of GMRES(30) with SPAI(0.3) having $\mathrm{nz}(M) = 21336$ (dashed line) and approximate factors algorithm (solid line). (b) The spectrum of the preconditioned matrix AWV_1^{-1}.

Matrix saylor4 I

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

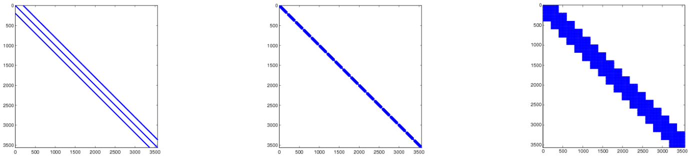
Let A to be the saylor4 matrix from Matrix Market arising in petroleum engineering simulation having $n = 3564$ with $\mathrm{nz}(A) = 22316$ and $\kappa(A) = 1E + 2$.

Use GMRES(20) to solve the preconditioned linear system (4).

Determine the sparsity pattern of $\mathcal{W}$ using numerical dropping with $\tau = 1e - 5$ and $p = 10$ and initial guess $W_0 = I$.

Choose $\mathcal{V}_1$ as the set of block diagonal matrices with blocksize $k = 198$. Use numerical dropping with $\tau = 1e - 5$ and $p = 20$.

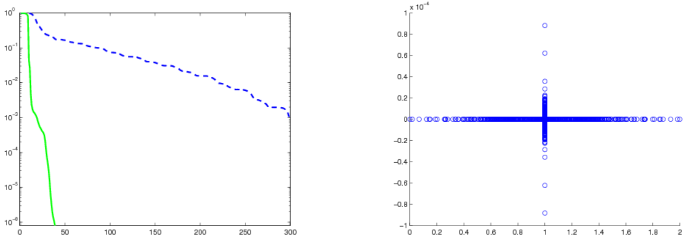
Matrix saylor4 II



(a) W (b) V_1 (c) WV_1^{-1}

Figure: Sparsity patterns of the factors W and V_1 and the preconditioner WV_1^{-1} after 5 iterations. Number of nonzeroes $\mathrm{nz}(W)$=26482, $\mathrm{nz}(V_1)$=15626 and $\mathrm{nz}(WV_1^{-1})$=2009304

Matrix `saylor4` III



(a) (b)

Figure: (a) Convergence of GMRES(20) with SPAI(0.1) using 10 refinements and 10 new non-zeroes per refinement having $\mathrm{nz}(M) = 202343$ (dashed line) and approximate factors algorithm (solid line). (b) The spectrum of the preconditioned matrix AWV_1^{-1}.

Nonsymmetric PDE I

Approximate factoring of the inverse
Mikko Byckling, Marko Huhtanen
Introduction
Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm
Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE
Conclusions

Consider the partial differential equation

$$-u_{xx} - u_{yy} + \tau(xu_x + yu_y) + \eta u = f, \qquad (5)$$

on unit square $\Omega = (0,1) \times (0,1)$ with Dirichlet boundary conditions on $\partial\Omega$. Choose $\tau = 10$ and $\eta = -100$.

Use centered differences with n gridpoints in both directions.

After discretization $A = H + V$, with $A \in \mathbb{R}^{n^2 \times n^2}$ where H and V correspond to x- and y-directions on the grid.

Nonsymmetric PDE II

Approximate factoring of the inverse

Mikko Byckling, Marko Huhtanen

Introduction

Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm

Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE

Conclusions

Let $\mathcal{W}$ be determined by the sparsity pattern of A^2 and $\mathcal{V}_1$ by the sparsity pattern of H^2 ($\mathcal{V}_1$ is block diagonal with pentadiagonal blocks of size n). Let initial guess $W_0 = I$.

Use FGMRES(30) preconditioned with GMRES(10)-SPAI or GMRES(10)-AF to solve the linear system (4).

$n \times n$	$\mathrm{nz}(A)$	$\mathrm{nz}(M)$	$\mathrm{nz}(W) + \mathrm{nz}(V_1)$	SPAI	AF
16×16	1216	7834	4196	16	19
32×32	4992	22513	17604	22	22
64×64	20224	83418	72068	51	30
128×128	81408	295121	291588	49	30
256×256	326656	1326396	1172996	121	64

Table: Iteration counts of FGMRES(30) for discretized (5) preconditioned with GMRES(10)-SPAI and GMRES(10)-AF.

Conclusions

Approximate factoring of the inverse

Mikko Byckling, Marko Huhtanen

Introduction

Approximate factors
Basic algorithm
Implementation
Computational complexity
Subspaces $\mathcal{W}$ and $\mathcal{V}_1$
Practical algorithm

Numerical experiments
Sherman2
Saylor4
Nonsymmetric PDE

Conclusions

- Approximate factors can be computed by using the power method to have preconditioners of the form WV_1^{-1}.
- To have preconditioners of good quality, about 5 to 10 iterations of the power method seems to suffice.
- Preconditioners can be dense and are nearly optimal.
- Choice of subspaces $\mathcal{W}$ and $\mathcal{V}_1$ is very important (and challenging) for obtaining a good preconditioner.
- Good sparsity structures for $\mathcal{W}$ and $\mathcal{V}_1$ can be computed by using approximate factors algorithm with numerical dropping.

M. Byckling and M. Huhtanen.
Approximate factoring of the inverse.
A submitted manuscript available at
http://math.tkk.fi/~mhuhtane/index.html.

M. Byckling.
PhD Thesis.
In preparation.

M. Grote and T. Huckle,
Parallel preconditioning with sparse approximate inverses.
SIAM J. Sci. Comput., 18:838–853, 1997.

NIST.
MatrixMarket
http://math.nist.gov/MatrixMarket/.

Y. Saad and M.H. Schultz.
GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems.
SIAM J. Sci. Stat. Comput., 7:856–869, 1986.

Y. Saad.
Iterative Methods for Sparse Linear Systems, 2nd edition.
SIAM, Philadelphia, PA, 2003.

Thank you for your interest!