



Preconditioners based on the ISM factorization*

Rafael Bru and Juana Cerdán and José Marín and José Mas and
Miroslav Tůma

Abstract

In this paper we survey our work on preconditioners based on the Inverse Sherman-Morrison factorization. The most important theoretical results are also summarized and some numerical conclusions are provided.

1 Introduction

The solution of linear systems

$$Ax = b, \quad (1)$$

where A is a nonsingular large and sparse matrix, is usually computed applying an iterative method, typically a Krylov method, see [17]. To get convergence, or to improve it, a preconditioner can be used. A preconditioner is a matrix M that approximates the inverse of A and can be constructed in several ways, even implicitly. When a preconditioner is used, instead of system (1) one of the systems

$$\begin{aligned} MAx &= Mb && \text{(left preconditioning)} \\ AMy &= b, \quad x = My && \text{(right preconditioning)} \\ M^{1/2}AM^{1/2}y &= M^{1/2}b \quad x = M^{1/2}y && \text{(symmetric preconditioning)} \end{aligned} \quad (2)$$

Key Words: Preconditioned iterative methods, sparse matrices, incomplete decompositions, approximate inverses, Sherman-Morrison formula.

2010 Mathematics Subject Classification: Primary 65F08, 65F10; Secondary 65F50.

Received: November, 2014.

Revised: November, 2014.

Accepted: February, 2015.

*This work was supported by the Spanish Ministerio de Economía y Competitividad under grant MTM2014-58159-P and by the project 13-06684S of the Grant Agency of the Czech Republic.

is solved by an iterative method.

There are many ways to compute a preconditioner. Two of them are of interest for us, the one that computes an approximation of A^{-1} , the so called *approximate inverse preconditioners*; the second is based on the computation of an *incomplete factorization of A* , typically the well known Incomplete LU (ILU) factorization or the Incomplete Cholesky (IC) factorization for the symmetric positive definite case.

Let us focus now on approximate inverse preconditioners, (see [3] for a detailed discussion). These preconditioners are applied by matrix-vector products in each iteration of the Krylov method. Then they are highly parallel. Some methods to obtain approximate inverse preconditioners are based on minimizing $\|I_n - AM\|_F$, subject to some sparsity constraints on M . The use of the Frobenius norm allows to reduce the problem to solving independent linear least square problems, which can be done also in parallel. Examples of this class are the SPAI [14] and MR [13] preconditioners. Other approaches compute the preconditioner M as a product of triangular factors by an A -biorthogonalization process, as the AINV method, see [1] and [2].

On the contrary, other preconditioners try to approximate A computing an incomplete LU [16] or Cholesky factorization. They are implicit because instead of computing the matrix M in (2), what it is computed is a factorization of its inverse, resulting in a factorization of a matrix that is close to the coefficient matrix A . Then, in each iteration instead of a matrix vector product with the matrix M two linear systems with triangular matrices must be solved. This implies that these kind of preconditioners are not as parallel as the approximate inverse ones. However they have some advantages. As A is usually much more sparse than its inverse, this implicit preconditioners can also be much more sparse than the explicit ones to get an equivalent improvement on the convergence of the iterative method.

To get sparse preconditioners it is necessary to annihilate entries during computation. A threshold is selected and entries are zeroed if their absolute values are less than this quantity, usually multiplied by some quantity as, for example, the norm of the corresponding row. That is, an entry is zeroed if it is small relative to some other relevant quantity. In addition a maximum number of entries can be allowed by row or column. This technique is the most usual when computing preconditioners of both types summarized previously.

The Inverse Sherman-Morrison (ISM) decomposition introduced in [8] that can be used to derive preconditioners of both types: approximate inverse preconditioners and incomplete LU or Cholesky preconditioners. In the next section we will review the general framework of the factorization and some theoretical results about the decomposition. In section 3 we will review preconditioners and we will also mention some other results. Our goal is to present

them in the order they have been known.

2 The ISM factorization

In this section we present the ISM factorization. The factorization was introduced in [8] and later developed in several works. During the process some changes in the original notation have been introduced to improve readability. In the sequel we will use the new notation.

The ISM factorization gives a shifted factorization of the inverse of a matrix and, as stated later, it also contains some of the LU factors of the matrix and its inverse. The starting point is the next result.

Theorem 1 ([8, Theorem 2.1 and Corollary 2.2]). *Let A and A_0 be two nonsingular matrices, and let $\{x_k\}_{k=1}^n$ and $\{y_k\}_{k=1}^n$ be two sets of vectors such that*

$$A = A_0 + \sum_{k=1}^n x_k y_k^T. \quad (3)$$

Suppose that $r_k = 1 + y_k^T A_{k-1}^{-1} x_k \neq 0$ for $k = 1, \dots, n$, where $A_{k-1} = A_0 + \sum_{i=1}^{k-1} x_i y_i^T$. Then A_k is nonsingular and $A_k^{-1} = A_{k-1}^{-1} - \frac{1}{r_k} A_{k-1}^{-1} x_k y_k^T A_{k-1}^{-1}$. Moreover the vectors

$$z_k := x_k - \sum_{i=1}^{k-1} \frac{v_i^T A_0^{-1} x_k}{r_i} z_i \quad (4)$$

and

$$v_k := y_k - \sum_{i=1}^{k-1} \frac{y_k^T A_0^{-1} z_i}{r_i} v_i \quad (5)$$

are well defined and satisfy the relations

$$\begin{aligned} A_{k-1}^{-1} x_k &= A_0^{-1} z_k \\ y_k^T A_{k-1}^{-1} &= v_k^T A_0^{-1} \end{aligned} \quad (6)$$

and

$$r_k = 1 + y_k^T A_0^{-1} z_k = 1 + v_k^T A_0^{-1} x_k \quad (7)$$

for $k = 1, \dots, n$. In addition

$$A_0^{-1} - A^{-1} = A_0^{-1} Z D V^T A_0^{-1}, \quad (8)$$

where $Z = [z_1, \dots, z_n]$, $D = \text{diag}(r_1, \dots, r_n)$ and $V = [v_1, \dots, v_n]$.

Note that conditions $r_k \neq 0$, $k = 1, \dots, n$ allow the use of the Sherman-Morrison theorem to prove nonsingularity of matrices A_k . But taking into account (6) and (7) the process can be rewritten without these matrices.

Different choices of A_0 , $\{x_k\}$ and $\{y_k\}$ vectors can lead to different factorizations of $A_0^{-1} - A^{-1}$. The choice that produces what we call the Inverse Sherman-Morrison factorization is

$$A_0 = sI, \quad x_k = e_k, \quad y_k = (a^k - se^k)^T, \quad (9)$$

where s is a positive parameter, e_k and e^k are the k th column and row respectively of the identity matrix, and a^k is the k th row of A . In the sequel we will denote the rows and columns of a matrix using superindexes and subindexes respectively.

With the choices (9), equations (4), (5) and (7) become

$$\begin{aligned} z_k &= e_k - \sum_{i=1}^{k-1} \frac{v_{ki}}{sr_i} z_i, \\ v_k &= y_k - \sum_{i=1}^{k-1} \frac{a^k z_i}{sr_i} v_i, \\ r_k &= 1 + \frac{y_k^T z_k}{s} = \frac{a^k z_k}{s} = 1 + \frac{v_{kk}}{s} \end{aligned} \quad (10)$$

Then, if $Z_s = [z_1, \dots, z_n]$, $D_s = \text{diag}(r_1, \dots, r_n)$ and $V_s = [v_1, \dots, v_n]$, the ISM factorization is given by

$$s^{-1}I - A^{-1} = s^{-2}Z_s D_s^{-1} V_s^T. \quad (11)$$

In [8, Lemma 3.1] the relation of the factors for different values of the parameter s are established, in particular it is proved that Z_s actually does not depend on s , so the subindex will be removed in the sequel. Algorithm 1 computes the ISM factorization of a given matrix A . It is worth to note the relation between vectors z_k and v_k .

The ISM factorization is strongly related with the LU factorization as was proved in [9].

Theorem 2. *A square matrix A has LDU factorization, $A = LDU$, if and only if it has ISM factorization $s^{-1}I - A^{-1} = s^{-2}ZD_s^{-1}V_s^T$.*

Then

$$D = s^{-1}D_s, \quad U = Z^{-1}, \quad V_s = U^T D - sL^{-T}.$$

Algorithm 1 Computation of the exact ISM factorization (11).

```

(1) for  $k = 1, \dots, n$ 
     $z_k = e_k$ 
     $v_k = (a^k - se^k)^T$ 
    for  $i = 1, \dots, k - 1$ 
         $z_k = z_k - \frac{v_{ki}}{sr_i} z_i$ 
         $v_k = v_k - \frac{a^k z_i}{sr_i} v_i$ 
    end for
     $r_k = 1 + v_{kk}/s$ 
end for
(2) Return  $Z = [z_1, z_2, \dots, z_n]$ ,  $V = [v_1, v_2, \dots, v_n]$  and
     $D = \text{diag}(r_1, r_2, \dots, r_n)$ .

```

3 Preconditioners based on the ISM factorization

In this section we review all preconditioners derived up to now from the ISM factorization. They are approximate inverse preconditioners and also ILU type preconditioners. We also mention the application of them to least squares problems.

3.1 Approximate inverse preconditioners

The exact process described in Algorithm 1 can be used to compute a preconditioner if a dropping strategy of small entries is implemented. In [8] the dropping strategy was based on the size of entries and several results about the incomplete factorization are proved. The most important one is that the incomplete process can be applied without breakdown to M-matrices. The approximate inverse preconditioner (AISM) was compared with AINV, showing similar robustness and performance.

A block extension of AISM (bAISM) was introduced in [12]. If A is block partitioned in such a way that diagonal blocks are square, the block ISM reads

$$A^{-1} = A_0^{-1} - A_0^{-1} U T^{-1} V^T A_0^{-1}, \quad (12)$$

where all matrices have the same block partition than A , U is block upper triangular and T is block diagonal.

Two interesting results were proved. If A_0 is taken as the identity matrix, then by [12, Lemma 3.2.] the k th diagonal block of T is exactly the k th pivot in the block LU factorization of A . Then the incomplete process does not breakdown if A is an H-matrix, [12, Theorem 3.4]. A similar result, [12, Theorem 3.7], is proved if A_0 is the diagonal block matrix formed with the

diagonal blocks of A . It is worth to mention that this second case reduces to the first one if a right block Jacobi scaling is previously done to A [12, Theorem 3.8].

3.2 ILU type preconditioners based on ISM

In [9] the relation between the ISM and LDU factorizations, that was previously stated for pivots, was extended in a fruitful way as stated in Theorem 2. From this result it follows that the ISM factorization contains the D (maybe scaled) and U factors of the LDU factorization of A , and the factors L^{-1} , U^{-1} of the LDU factorization of A^{-1} , the other factor D^{-1} is easily computed as sD_s^{-1} . This situation is specially interesting when A is a symmetric positive definite matrix. Then all factors of the LDL^T factorizations of A are available and it is possible to get a preconditioner computing an incomplete Cholesky factorization. Moreover, since all factors of the Cholesky factorization of A^{-1} are also computed, special dropping strategies can be used, namely those based on the results obtained by M. Bollhöfer and Y. Saad for LU factorizations, [6, 7]. This robust strategy drops entries by value, but the value is compared with the norm of a row of the inverse. That is, given a drop tolerance τ an entry of the incomplete factor \hat{L} is zeroed if

$$|\hat{l}_{jk}| \|e^k \hat{L}^{-1}\| \leq \tau.$$

Also the dual dropping criteria can be used. If $\hat{\ell}_{jk}$ denotes the entry of j th row, k th column of the incomplete factor \hat{L}^{-1} , it is zeroed if

$$|\hat{\ell}_{jk}| \|e^k \hat{L}\| \leq \tau.$$

There is a balance between entries of the matrices \hat{L} and \hat{L}^{-1} computed by the preconditioner obtained implementing this dual dropping strategies to sparsify matrices in the ISM process given in Algorithm 1. So it is called Balanced Incomplete Factorization (BIF). This balance gives raise to robust preconditioners as numerical experiments presented in [9] confirm. In addition they also show that the BIF preconditioner has a very good behavior from the point of view of the selection of the dropping parameter. As it is well known as the value of the dropping parameter decreases the density of the preconditioner increases. Although theoretically the number of iterations should decrease, for some methods and matrices this is not true. Even, in some cases, the preconditioner stops being able to get convergence for some interval of the dropping parameter, and then the method converges again. This is not the case of BIF, the method is robust and the number of iterations varies smoothly with a decreasing trend, see Figures 5.1–5.3 in [9]. On the other hand we get

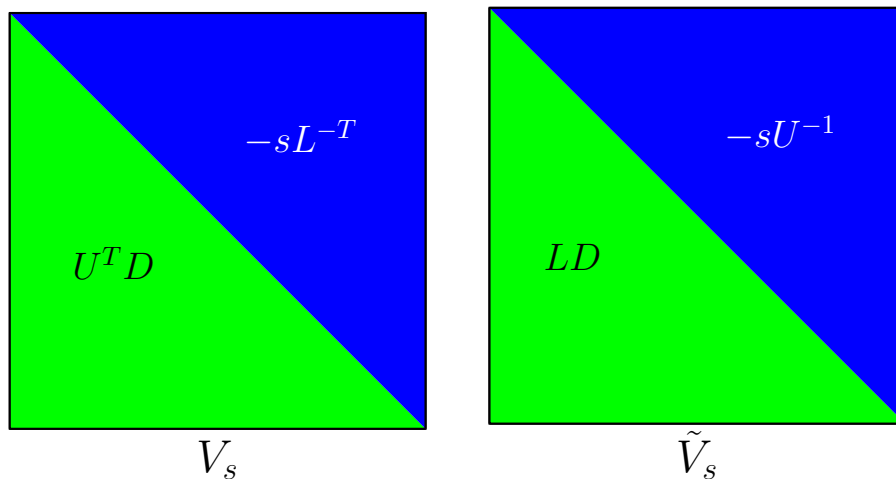


Figure 1: Factors of the LDU factorizations of A and A^{-1} included in V_s and \tilde{V}_s . Green parts are factors of A , and blue parts are factors of A^{-1} .

good results with very sparse preconditioners. Let us mention at this point that all codes we use to compute preconditioners implement also a restriction in the number of nonzero entries of the matrix, which probably is the reason that explain why we can not get denser preconditioners.

3.3 ISM preconditioners for general nonsingular matrices: nBIF

As mentioned above, when A is a nonsymmetric matrix from the ISM factorization one can retrieve all factors in $A^{-1} = U^{-1}D^{-1}L^{-1}$, and factors D and U in $A = LDU$. To get the missed factor L it is proposed in [10] to compute the ISM factorization of A^T .

To fix notation let

$$s^{-1}I - A^{-T} = s^{-2}\tilde{Z}\tilde{D}_s^{-1}\tilde{V}_s^T,$$

the ISM factorization of the inverse of A^T . Then

$$V_s = U^T D - sL^{-T} \quad \text{and} \quad \tilde{V}_s = LD - sU^{-1},$$

as depicted in Figure 1. As a consequence Z and \tilde{Z} are no longer needed.

Moreover, as stated in Lemma 3.1 of [10]

$$v_{pk} = s \frac{\tilde{v}_{kp}}{d_k} - \sum_{i=p+1}^{k-1} \frac{\tilde{v}_{ki}}{d_i} v_{pi}. \quad \text{for } p < k,$$

A dual expression allows computation of \tilde{v}_{pk} , for $p < k$, and then the strict upper triangular parts of V_s and \tilde{V}_s are computed in a new way. Entries of factors of the inverse are used to compute entries of factors of the matrix and vice versa. In BIF they were used only to drop entries.

Numerical results in [10] show the same that was already true for BIF. The method is robust and the dependence of the dropping tolerance is not critical.

3.4 BIF for least squares problems

It is clear that BIF can be applied to the iterative solution of least squares problems

$$\min_x \|b - Ax\|_2$$

when the coefficient matrix has full column rank (and hence has at least as many rows as columns). If this is the case the matrix $A^T A$ is symmetric and positive definite and a BIF preconditioner can be computed.

Anyway some particular aspects of the least squares problems should be taken into account. Since $A^T A$ is much denser than A , to get a sparse preconditioner a lot of entries should be dropped, and some instabilities may appear. To prevent this possibility two additional strategies are used in [11] to improve robustness. The first one is to replace the way that the pivots are computed. Instead of to compute them as in (10), that is as sr_i , they are computed as $z_i^T A^T A z_i$. The second one follows the ideas of Tismenetsky in [18], by storing additional entries that are used only in computations, but finally discarded a better preconditioner is computed, see Algorithm 2.2 in [11]. We refer to this preconditioner as lsBIF.

In [11] lsBIF was compared with IC, an incomplete Cholesky factorization preconditioner see [4] for implementation details; AINV; and Cholesky incomplete modified Gram Schmidt, CIMGS [19], see [15] for details of the implementation. All preconditioners were used to accelerate the conjugate gradient for least squares (CGLS) [5] iterative method. Results show the robustness of the method and also its stability with respect the choice of the threshold.

4 Conclusions

The ISM factorization can be used to compute robust, sparse and efficient preconditioners, taking advantage of the fact that it computes factors of the LDU factorization of the coefficient matrix and also of its inverse. Numerical experiments exhibit good performance for all preconditioners developed: direct ones, as BIF for symmetric matrices, nBIF for general matrices and lsBIF for least squares problems with full column rank; as well as those of approximate inverse type (AISM and bAISM).

The relation between the factorization and the LDU decomposition also means that the incomplete factorization can be computed breakdown free for H-matrices, and in the case of full column rank matrices for the normal equations.

References

- [1] M. Benzi, C. D. Meyer, and M. Tũma. A sparse approximate inverse preconditioner for the conjugate gradient method. *SIAM J. Sci. Comput.*, 17(5):1135–1149, 1996.
- [2] M. Benzi and M. Tũma. A sparse approximate inverse preconditioner for nonsymmetric linear systems. *SIAM J. Sci. Comput.*, 19(3):968–994, 1998.
- [3] M. Benzi and M. Tũma. A comparative study of sparse approximate inverse preconditioners. *Appl. Numer. Math.*, 30(2-3):305–340, 1999.
- [4] M. Benzi and M. Tũma. A robust preconditioner with low memory requirements for large sparse least squares problems. *SIAM J. Sci. Comput.*, 25(2):499–512, 2003.
- [5] Å. Björck. *Numerical methods for Least Squares Problems*. SIAM, Philadelphia, 1996.
- [6] M. Bollhöfer. A robust and efficient *ILU* that incorporates the growth of the inverse triangular factors. *SIAM J. Sci. Comput.*, 25(1):86–103, 2003.
- [7] M. Bollhöfer and Y. Saad. On the relations between *ILUs* and factored approximate inverses. *SIAM J. Matrix Anal. Appl.*, 24(1):219–237, 2002.
- [8] R. Bru, J. Cerdán, J. Marín, and J. Mas. Preconditioning sparse nonsymmetric linear systems with the Sherman-Morrison formula. *SIAM J. Sci. Comput.*, 25(2):701–715, 2003.

- [9] R. Bru, J. Marín, J. Mas, and M. Tuma. Balanced incomplete factorization. *SIAM J. Sci. Comput.*, 30(5):2302–2318, 2008.
- [10] R. Bru, J. Marín, J. Mas, and M. Tuma. Improved balanced incomplete factorization. *SIAM J. Matrix Anal. Appl.*, 31(5):2431–2452, 2010.
- [11] R. Bru, J. Marín, J. Mas, and M. Tuma. Preconditioned iterative methods for solving linear least squares problems. *SIAM J. Sci. Comput.*, 36(4):A2002–A2022, 2014.
- [12] J. Cerdán, T. Faraj, N. Malla, J. Marín, and J Mas. Block approximate inverse preconditioners for sparse nonsymmetric linear systems. *ETNA*, 37:23–40, 2010.
- [13] E. Chow and Y. Saad. Approximate inverse preconditioners via sparse-sparse iterations. *SIAM J. Sci. Comput.*, 19(3):995–1023, 1998.
- [14] M. J. Grote and T. Huckle. Parallel preconditioning with sparse approximate inverses. *SIAM J. Sci. Comput.*, 18(3):838–853, 1997.
- [15] I. E. Kaporin. High quality preconditioning of a general symmetric positive definite matrix based on its $U^T U + U^T R + R^T U$ decomposition. *Numer. Linear Algebra Appl.*, 5:483–509, 1998.
- [16] Y. Saad. ILUT: a dual threshold incomplete LU factorization. *Numer. Linear Algebra Appl.*, 1(4):387–402, 1994.
- [17] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Co., Boston, 1996.
- [18] M. Tismenetsky. A new preconditioning technique for solving large sparse linear systems. *Linear Algebra Appl.*, 154–156:331–353, 1991.
- [19] X. Wang, K. A. Gallivan, and R. Bramley. CIMGS: an incomplete orthogonal factorization preconditioner. *SIAM J. Sci. Comput.*, 18(2):516–536, 1997.

Rafael Bru,
Instituto de Matemática Multidisciplinar,
Universitat Politècnica de Valencia,
Camí de Vera s/n, 46022 Valencia, Spain.
Email: rbru@imm.upv.es

Juana Cerdán,
Instituto de Matemática Multidisciplinar,
Universitat Politècnica de Valencia,
Camí de Vera s/n, 46022 Valencia, Spain.
Email: jcerdan@imm.upv.es

José Marín,
Instituto de Matemática Multidisciplinar,
Universitat Politècnica de Valencia,
Camí de Vera s/n, 46022 Valencia, Spain.
Email: jmarinma@imm.upv.es

José Mas,
Instituto de Matemática Multidisciplinar,
Universitat Politècnica de Valencia,
Camí de Vera s/n, 46022 Valencia, Spain.
Email: jmasm@imm.upv.es

Miroslav Tůma,
Institute of Computer Science,
Academy of Sciences of the Czech Republic,
Pod vodárenskouvěží 2, 182 07 Prague 8, Czech Republic.
Email: tuma@cs.cas.cz

