Preconditioning with the Sherman-Morrison formula

R. Bru J. Cerdán J. Marín J. Mas

Departament de Matemàtica Aplicada
Universitat Politècnica de València
Camí de Vera, 14, 46022 València (Spain)
{rbru,jcerdan,jmarinma,jmasm}@mat.upv.es

Abstract

Let

$$Ax = b \tag{1}$$

be a linear system of n equations with n unknowns, where A is a large and sparse matrix. An approximate solution of (1) is usually obtained by using preconditioned iterative krylov-type methods.

Several preconditioning techniques can be applied. Among them, we are interested in the study of approximate inverse preconditioners. With such preconditioners, an approximation of the inverse of A, namely $M^{-1} \approx A^{-1}$, is computed and stored explicitly. Thus, the preconditioning step can be done by computing matrix-vector products.

In this work, we show how the matrix $A^{-1} - I_n$, where I_n is the $n \times n$ identity matrix, can be used as preconditioner. For that the matrix $A^{-1} - I_n$ is factorized in the form $U\Omega V^T$ using the Sherman-Morrison formula. If the factorization process is done incompletely by applying a dropping strategy during the computation of the factors U and V, an approximate factorization for the matrix $A^{-1} - I_n$ is computed.

The following well known result from Sherman and Morrison will be used to invert a matrix in n steps.

Let B be a $n \times n$ nonsingular matrix, and let x and y be two vectors in $\mathbf{R}^n$ such that $r = 1 + y^T B^{-1} x \neq 0$, then the matrix $A = B + xy^T$ is nonsingular, and its inverse is $A^{-1} = B^{-1} - r^{-1} B^{-1} x y^T B^{-1}$.

Let $\{x_k\}_{k=1}^n$ and $\{y_k\}_{k=1}^n$ be two sets of vectors in $\mathbf{R}^n$, and let A_0 be a nonsingular $n \times n$ matrix whose inverse is either known or easy to obtain, and such that

$$A = A_0 + \sum_{k=1}^{n} x_k y_k^T \,. \tag{2}$$

We define $A_k := A_0 + \sum_{i=1}^k x_i y_i^T$ with $k = 1, \ldots, n$. Then

$$\left\{ \begin{array}{l} A_{k+1} = A_k + x_{k+1} y_{k+1}^T \\ A_n = A \end{array} \right. .$$

If x_k and y_k are vectors such that $r_k = 1 + y_k^T A_{k-1}^{-1} x_k \neq 0$, then

$$A_k^{-1} = A_{k-1}^{-1} - \frac{1}{r_k} A_{k-1}^{-1} x_k y_k^T A_{k-1}^{-1}$$

for $k = 1, \ldots, n$, with $A_n^{-1} = A^{-1}$.

This process can be written as follows

$$A^{-1} = A_0^{-1} - \sum_{k=1}^{n} \frac{1}{r_k} A_{k-1}^{-1} x_k y_k^T A_{k-1}^{-1}$$

or

$$A_0^{-1} - A^{-1} = \sum_{k=1}^{n} \frac{1}{r_k} A_{k-1}^{-1} x_k \,. y_k^T A_{k-1}^{-1}$$

In matrix notation

$$A_0^{-1} - A^{-1} = \Phi \Omega^{-1} \Psi^T \,,$$

where

$$\Phi = \left[\begin{array}{cccc} A_0^{-1} x_1 & A_1^{-1} x_2 & \cdots & A_{n-1}^{-1} x_n \end{array} \right],$$

$$\Omega^{-1} = \left[\begin{array}{cccc} r_1^{-1} & & & \\ & r_2^{-1} & & \\ & & \ddots & \\ & & & r_n^{-1} \end{array} \right], \quad \text{and} \quad \Psi^T = \left[\begin{array}{c} y_1^T A_0^{-1} \\ y_2^T A_1^{-1} \\ \vdots \\ y_n^T A_{n-1}^{-1} \end{array} \right].$$

This process depends explicitly on the matrices $\{A_k\}_{k=1}^n$. The following results show how this factorization can be computed without these matrices explicitly appearing.

Theorem 1 *Let A and A_0 be two nonsingular matrices, and let $\{x_k\}_{k=1}^n$ and $\{y_k\}_{k=1}^n$ be two set of vectors such that condition (2) is satisfied. Moreover, suppose that $r_k = 1 + y_k^T A_{k-1}^{-1} x_k \neq 0$ for $k = 1, \ldots, n$, where $A_{k-1} = A_0 + \sum_{i=1}^{k-1} x_i y_i^T$. Then*

$$u_k := x_k - \sum_{i=1}^{k-1} \frac{v_i^T A_0^{-1} x_k}{r_i} u_i \tag{3}$$

and

$$v_k := y_k - \sum_{i=1}^{k-1} \frac{y_k^T A_0^{-1} u_i}{r_i} v_i \tag{4}$$

are well defined for $k = 1, \ldots, n$. In addition, the relations

$$\begin{array}{rcl} A_{k-1}^{-1} x_k & = & A_0^{-1} u_k \\ y_k^T A_{k-1}^{-1} & = & v_k^T A_0^{-1} \end{array}$$

and $r_k = 1 + y_k^T A_0^{-1} u_k = 1 + v_k^T A_0^{-1} x_k$ are satisfied.

Corollary 1 *Under the assumptions of Theorem 1, let* $U := \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$ *and* $V := \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$, *where* u_k *and* v_k *are defined in (3) and (4), respectively. Then*

$$A_0^{-1} - A^{-1} = A_0^{-1} U \Omega^{-1} V^T A_0^{-1} \tag{5}$$

where $\Omega = \mathrm{diag}(r_1, r_2, \ldots, r_n)$.

Corollary 2 *Under the assumptions of Theorem 1 and Corollary 1 there exists unique unit upper triangular matrices* T_x *and* T_y *such that*

$$X = UT_X \qquad Y = VT_Y$$

where X *and* Y *are matrices whose columns are the vectors* x_i *and* y_i, *respectively,*

$$T_X = \begin{bmatrix} 1 & t_{12} & & t_{1n} \\ 0 & 1 & & t_{2n} \\ \vdots & & \ddots & \\ & & & t_{n-1,n} \\ 0 & & \cdots & 1 \end{bmatrix}, \qquad t_{ij} = \frac{v_i^T A_0^{-1} x_j}{r_i}, \quad i < j,$$

and

$$T_Y = \begin{bmatrix} 1 & \bar{t}_{12} & & \bar{t}_{1n} \\ 0 & 1 & & \bar{t}_{2n} \\ \vdots & & \ddots & \\ & & & \bar{t}_{n-1,n} \\ 0 & & \cdots & 1 \end{bmatrix}, \qquad \bar{t}_{ij} = \frac{y_j^T A_0^{-1} u_i}{r_i}, \quad i < j,$$

Different choices of the vectors $\{x_k\}$ and $\{y_k\}$ and the matrix A_0 will lead to different factorizations of the matrix $A_0^{-1} - A^{-1}$. In particular, in this work we take

$$\begin{aligned} A_0 &= I_n\,, \\ x_k &= e_k\,, \\ y_k &= (a^k)^T - x_k\,, \end{aligned}$$

where a^k denotes the kth row of the matrix A and e_k is the kth vector of the canonical basis in $\mathbf{R}^n$. It follows that the factorization (5) simplifies to

$$U\Omega^{-1}V^T = I_n - A^{-1} \tag{6}$$

We note that, since $x_k = e_k$, the matrix X in Corollary 2 is the identity matrix. Therefore $U = T_X^{-1}$ is a nonsingular unit upper triangular matrix.

Denoting with $(x)_k$ the kth element of a vector x, the basic algorithm to compute the factorization (6) can be written as follows.

Even if A is a sparse matrix, in general the matrices U and V in equation (6) tend to be dense. If a dropping strategy is applied, the factorization algorithm computes sparse matrices $\bar{U}$ and $\bar{V}$, and a diagonal matrix $\bar{\Omega}$ such that

$$\bar{U}\bar{\Omega}^{-1}\bar{V}^T \approx I_n - A^{-1}, \tag{7}$$

BASIC FACTORIZATION ALGORITHM

(1) **Set** $x_k = e_k$, $y_k = (a^k)^T - x_k$, $(k = 1, \dots, n)$
(2) **for** $k = 1, \dots, n$
$\quad u_k = x_k$
$\quad v_k = y_k$
$\quad$ **for** $i = 1, \dots, k-1$
$\quad\quad u_k = u_k - \frac{(v_i^T)_k}{r_i} u_i$
$\quad\quad v_k = v_k - \frac{y_k^T u_i}{r_i} v_i$
$\quad$ **end for**
$\quad r_k = 1 + (v_k^T)_k$
end for
(3) **Return** $U = [u_1, u_2, \cdots, u_n]$, $V = [v_1, v_2, \cdots, v_n]$ and
$\Omega = \mathrm{diag}(r_1, r_2, \dots, r_n)$.

Figure 1: Computation of the factorization (6).

or, equivalently

$$M^{-1} := I_n - \bar{U}\bar{\Omega}^{-1}\bar{V}^T \approx A^{-1}. \tag{8}$$

Equation (7) shows the expression of the explicit preconditioner which will be used for preconditioning different Krylov-type iterative methods.

The results of the numerical experiments obtained for a representative set of matrices will be presented.