A note on the pivoted symmetric LR iteration

K. Veselić *

Abstract

We prove that the diagonally pivoted symmetric LR algorithm on a positive definite matrix is globally convergent.

The "symmetric" or "Cholesky" LR iteration is a fairly old method of eigenreduction of a positive definite Hermitian matrix H. It reads

$$\begin{array}{rcccl} H = H_0 & = & R_0^* R_0 & & \\ H_1 & = & R_0 R_0^* & = & R_1^* R_1 \\ \vdots & & & & \end{array} \tag{1}$$

This process is linearly convergent [6], [5]. Recently, its singular value 'implicit' equivalent

$$R_k^* = Q_k R_{k+1},\ Q_k \text{ unitary},\ R_k \text{ upper triangular}, \tag{2}$$

was studied in [3].

An obvious modification of the algorithm (2) includes pivoting. Thus modified, (2) reads

$$R_k^* = Q_k R_{k+1} P_k, \tag{3}$$

where P_k is a permutation of standard column pivoting or, equivalently, of the diagonal pivoting within the 'explicit' algorithm (1). This means that $R = R_k$ has the property

$$r_{ii} \geq \sqrt{|r_{il}|^2 + \cdots + |r_{ll}|^2},\ i = 1, \ldots, n,\ l \geq i. \tag{4}$$

Although the practical use of pivoting is limited to first few steps, it is of interest to know whether the pivoted algorithm itself is globally convergent. The answer is affirmative and this will be shown in our main theorem below.

The importance of pivoting was stressed in [2], Cor. 5.4, 5.5, where it was shown that the pivoted implicit Cholesky iteration computes the singular values with high relative accuracy. We produce a simple example which shows that the non-pivoted algorithm really is worse in this respect. Set

$$R = \begin{pmatrix} 1e-12 & 2e-12 \\ 1 & 1 \end{pmatrix}$$

Here the pivoted algorithm gives the correct small singular value $7.07106781186548e-13$ as expected, while the non-pivoted algorithm gives only $7.07099414738691e-13$. (Both algorithms were implemented in MATLAB by using its routine qr.[1])

*Lehrgebiet Mathematische Physik, Fernuniversität Hagen, Postf. 940, 58084 Hagen Germany, email: kresimir.veselic@fernuni-hagen.de

[1]The genuine MATLAB routine svd was also unsatisfactory, its outputs were $min(svd(R)) = 7.07012634145304e-13$ and $min(svd(R')) = 7.07106781207136e-13$.

Theorem 1 *Let $R = R_0$ be a non-singular matrix and let the sequence R_k be given by (3) and (4). Then R_k converges to a diagonal matrix, whose diagonals are the singular values of R_0. The convergence is linear.*

Proof. We will analyse a typical step of the algorithm. We set $R_k = R'$ and $R_{k+1} = R$.

Lemma 2 *Let*

$$R'^* = QRP, \tag{5}$$

with Q unitary and P a permutation such that R satisfies (4).[2] *Then*

$$\prod_{i=1}^{n}(R'R'^*)_{ii} \geq \prod_{i=1}^{n}(RR^*)_{ii} + \sum_{l=1}^{n-1}\sum_{j=l+1}^{n}|r_{li}|^2\sum_{j=l+1}^{i}|r_{ji}|^2\sigma_l^{2n-2l-2} \tag{6}$$

where σ_i is the smallest singular value of the matrix

$$\begin{pmatrix} r_{ii} & \cdots & r_{in} \\ & \ddots & \vdots \\ & & r_{nn} \end{pmatrix} \tag{7}$$

Proof of the Lemma. By (5) we have

$$\prod_{i=1}^{n}(R'R'^*)_{ii} = \prod_{i=1}^{n}(P^*R^*RP)_{ii} = \prod_{i=1}^{n}(R^*R)_{ii}$$

so the whole analysis takes place on R. Our next step is, in fact, a revisiting and a sharpening of an argument from [1].

By (4) we have

$$r_{11}^2(|r_{12}|^2 + r_{22}^2) = (r_{11}^2 + |r_{12}|^2)r_{22}^2 + |r_{12}|^2(|r_{12}|^2 + r_{11}^2 - r_{22}^2 - |r_{12}|^2) \geq |r_{12}|^4,$$

$$(r_{11}^2 + |r_{12}|^2)(|r_{13}|^2 + |r_{23}|^2 + r_{33}^2) =$$
$$(r_{11}^2 + |r_{12}|^2 + |r_{13}|^2)(|r_{23}|^2 + r_{33}^2) + |r_{13}|^2(|r_{13}|^2 + r_{11}^2 - r_{33}^2)$$
$$\geq (r_{11}^2 + |r_{12}|^2 + |r_{13}|^2)(|r_{23}|^2 + r_{33}^2) + |r_{13}|^2(|r_{12}|^2 + |r_{13}|^2)$$

and finally

$$(r_{11}^2 + \cdots + |r_{1n-1}|^2)(|r_{1n}|^2 + \cdots + r_{nn}^2) \geq$$
$$(r_{11}^2 + \cdots + |r_{1n}|^2)(|r_{2n}|^2 + \cdots + r_{nn}^2) + |r_{1n}|^2(|r_{1n}|^2 + \cdots + |r_{1n}|^2)$$

Using the fact that the Euclidian norm of any row or column in R is bounded from below by the smallest singular value σ_1 we obtain

$$\prod_{i=1}^{n}(R^*R)_{ii} \geq (r_{11}^2 + \cdots + |r_{1n-1}|^2)r_{22}^2(|r_{23}|^2 + r_{33}^2)\cdots(|r_{2n}|^2 + \cdots + r_{nn}^2)$$

$$+\sum_{i=2}^{n}|r_{1i}|^2\sum_{j=2}^{i}|r_{1j}|^2\sigma_1^{2n-4}$$

This can be continued to give (6). This proves the lemma.

[2] R' itself need not be triangular.

We continue with the proof of the theorem. Introducing again the sequence index k (6) can be written as

$$\prod_{i=1}^{n}(R'_k R'^*_k)_{ii} \geq \prod_{i=1}^{n}(R_{k+1}R^*_{k+1})_{ii} + \sum_{l=1}^{n-1}\sum_{j=l+1}^{n}|r_{li}^{(k+1)}|^2 \sum_{j=l+1}^{i}|r_{ji}^{(k+1)}|^2 \sigma_{lk}^{2n-2l-2} \tag{8}$$

All monomials within the triple sum are non-negative, thus the diagonal products are convergent and the triple sum itself tends to zero. If σ_{lk} were independent of k our theorem would already be proved. So, we have to use induction once more: the values σ_{1k}, are the singular values of R_k and they do not depend on k. Hence by (8)

$$r_{12}^{(k}, \ldots, r_{1n}^{(k)} \to 0$$

Now the continuity property of the singular values implies that σ_{2k} tends to a singular value of R_0 and thus stays bounded away from zero. Again by (8),

$$r_{23}^{(k)}, \ldots, r_{2n}^{(k)} \to 0$$

By induction, all off-diagonals of R_k tend to zero. Now by (4) the ordering

$$r_{11}^{(k)} \geq \cdots \geq r_{nn}^{(k)} \to 0$$

is preserved during the process, so the diagonal elements are themselves convergent to the singular values of R_0.

Concerning the speed of the convergence, we can immediately see that possible groups of multiple singular values will be approached by adjacent diagonal elements due to the ordering. So, for k large enough the pivoting will be performed at most within corresponding diagonal blocks. Now the existing convergence theory ([6]) insures that the off-diagonal blocks (punctured positions on the picture below)

$$\begin{array}{cccccc} * & * & \cdot & \cdot & \cdot & \cdot \\ & * & \cdot & \cdot & \cdot & \cdot \\ & & * & \cdot & \cdot & \cdot \\ & & & * & * & * \\ & & & & * & * \\ & & & & & * \end{array} \tag{9}$$

tend linearly to zero. The rest of the off-diagonal (starred positions) is even quadratically small, according to [4]. Q.E.D.

References

[1] J. Demmel and K. Veselić, Jacobi's method is more accurate than QR. SIAM J. Matr. Anal. Appl. **13** 1204-1246 (1992)

[2] Z. Drmač, Computing the singular and the generalized singular values, Ph.D. thesis, Fernuniversität Hagen 1994.

[3] K. Fernando and B. Parlett, Implicit Cholesky algorithms for singular values and vectors, Center for Pure and Applied Mathematics UC Berkeley 1993.

[4] V. Hari, On the quadratic convergence of the serial SVD Jacobi methods for triangular matrices, SIAM J. Sci. Stat. Comp. **10** 1076-1096 (1989)

[5] H. Rutishauser, Vorlesungen über Numerische Mathematik, Birkhäuser, Basel 1976.

[6] J. H. Wilkinson, Algebraic eigenvalue problem, Oxford University Press, Oxford, 1965.