Journal of Computational and Applied Mathematics 19 (1987) 293–299
North-Holland

Detection of near-singularity in Cholesky and LDL^T factorizations

Per Christian HANSEN *
Copenhagen University Observatory, Øster Voldgade 3, DK-1350 Copenhagen K, Denmark

Received 4 February 1986
Revised 4 February 1987

Abstract: In sparse matrix applications it is often important to implement the Cholesky and LDL^T factorization methods without pivoting in order to avoid excess fillin. We consider methods for detection of a nearly singular matrix by means of these factorizations without pivoting and demonstrate that a technique based on estimation of the smallest eigenvalue via inverse iteration will always reveal a nearly singular matrix.

Keywords: Nearly singular matrices, Cholesky and LDL^T factorization, inverse iteration.

1. Introduction

It is well-known that a singular or nearly singular symmetric nonnegative definite matrix can always be detected during a Cholesky or LDL^T factorization when these methods are implemented with complete symmetric pivoting, such as in e.g. the LINPACK program SCHDC [3]. In sparse matrix applications, however, it is usually necessary to use a pre-selected row ordering or to use pivoting for sparsity in order to avoid excess fillins during the factorization, which means that the security associated with the complete pivoting is lost. In such cases it is therefore necessary to find an alternative reliable strategy to detect a singular or nearly singular matrix.

In this paper we consider four alternative methods for detection of near-singularity in Cholesky and LDL^T factorizations. The requirement to these methods is that they should require little computational overhead to the factorization process. We demonstrate that a method based on estimation of the smallest eigenvalue of the matrix, such as used in the LINPACK condition estimator SPOCO [2,3], is a suitable method since it will always reveal a nearly singular matrix in only $O(n^2)$ flops.

We stress that in this presentation we are only concerned with *detection* of a nearly singular matrix, and not in methods for solving problems with such matrices.

* This work was supported by the Danish Space Board.

2. The Cholesky and LDL^T factorizations

In this representation, the formulations from [7] are used. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. Then the *Cholesky* and LDL^T *factorizations* are given by:

$$A = CC^T = LDL^T. \tag{1}$$

Here, the 'Cholesky triangle' $C \in \mathbb{R}^{n\times n}$ is lower triangular with positive diagonal elements, $L \in \mathbb{R}^{n\times n}$ is unit lower triangular, and $D \in \mathbb{R}^{n\times n}$ is diagonal with positive diagonal elements. The factorizations are computed by the following algorithms:

$$\left.\begin{aligned} c_{kk} &= \left[a_{kk} - \sum_{j=1}^{k-1} c_{kj}^2\right]^{1/2}, \\ c_{ik} &= \frac{1}{c_{kk}}\left[a_{ik} - \sum_{j=1}^{k-1} c_{ij}c_{kj}\right], \quad i = k+1,\ldots,n, \end{aligned}\right\} \quad k = 1, 2, \ldots, n; \tag{2}$$

$$\left.\begin{aligned} l_{kj} &= \tilde{l}_{kj}/d_{jj}, \quad j = 1, 2, \ldots, k-1, \\ d_{kk} &= a_{kk} - \sum_{j=1}^{k-1} \tilde{l}_{kj}l_{kj}, \\ \tilde{l}_{ik} &= a_{ik} - \sum_{j=1}^{k-1} \tilde{l}_{ij}l_{kj}, \quad i = k+1,\ldots,n, \end{aligned}\right\} \quad k = 1, 2, \ldots, n. \tag{3}$$

Let $\bar{C}$, $\bar{L}$, and $\bar{D}$ denote the matrices computed by means of the above algorithms (2) and (3). It is well-known that these matrices constitute close factorizations of A in the sense that [5, §5.2]:

$$A + E^{(1)} = \bar{C}\bar{C}^T, \qquad A + E^{(2)} = \bar{L}\bar{D}\bar{L}^T, \tag{4a, b}$$

where $E^{(1)}$ and $E^{(2)}$ are perturbations due to rounding errors during the factorization process. It can be shown that the norms of $E^{(1)}$ and $E^{(2)}$ are bounded by:

$$\| E^{(1)} \|_2 \leq c_n^{(1)}\epsilon_M \| A \|_2, \qquad \| E^{(2)} \|_2 \leq c_n^{(2)}\epsilon_M \| A \|_2 \tag{5a,b}$$

where ϵ_M is the machine precision, and $c_n^{(1)}$ and $c_n^{(2)}$ are small constants depending on n. Meinguet [8] has recently given an expression for $c_n^{(1)}$ which, for large n, can be approximated as:

$$c_n^{(1)} \approx \tfrac{2}{3}n^{3/2}. \tag{6}$$

A similar rounding error analysis of the LDL^T factorization has not been performed, but due to the similarities of the two methods we shall assume that $c_n^{(2)}$ can be approximated by the same expression (6). These results show that both factorizations are numerically stable also in absence of pivoting.

When the matrix A is singular or nearly singular, the above algorithms without pivoting are not guaranteed to reveal the near-singularity of A by one or more small diagonal elements in the computed $\bar{C}$ or $\bar{D}$ matrices. This can, however, be achieved if the algorithm is implemented with pivoting [3, §8.1]. The pivoting must obviously be symmetric, and a reasonable pivoting strategy is to select the pivot element in the kth step of (2) and (3) such that c_{kk} or d_{kk} becomes as large

as possible (similar to the column pivoting strategy in Q-R factorization). A 'look ahead' strategy is therefore required to select the pivot row and column p such that

$$h_p \geqslant h_i, \quad i = k, k+1, \ldots, n \tag{7}$$

where the candidate pivots are

$$h_i = a_{ii} - \sum_{j=1}^{k-1} c_{ij}^2 \quad \text{(Cholesky)}, \tag{8a}$$

$$h_i = a_{ii} - \sum_{j=1}^{k-1} \bar{l}_{ij}^2 / d_{ij} \quad (LDL^{\mathrm{T}}). \tag{8b}$$

If this pivoting strategy is used, a nearly singular matrix is *guaranteed* to be revealed by at least one computed diagonal element of $\bar{C}$ or $\bar{D}$ being small compared to the corresponding diagonal element of A. The computational overhead of this strategy is $O(n^2)$ flops and the strategy is therefore suitable for dense matrices.

3. Strategies that reduce pivoting

When A is large and sparse, complete pivoting may lead to a considerable amount of fillin, thus destroying the sparsity of A. It is therefore required to introduce a strategy that keeps pivoting at a minimum and which is guaranteed to detect a nearly singular matrix with only a small computational overhead. One can either try to design an alternative method that produces one or more small diagonal elements in $\bar{C}$ and $\bar{D}$, or use a method that detects the near-singularity of A by estimating its smallest eigenvalue from $\bar{C}$ or $\bar{L}$, $\bar{D}$.

We consider the following three strategies for producing small diagonal elements in $\bar{C}$ and $\bar{D}$:

(1) perform all accumulations in double precision and do not perform any pivoting;

(2) allow pivoting only if the maximum pivot h_p is considerably larger that the present pivot h_k, i.e., when

$$\tau h_p > h_k \tag{9}$$

where $0 < \tau < 1$ is a parameter to be chosen;

(3) allow pivoting in the final n_{piv} steps only; i.e., perform pivoting when

$$k \in \{ n - n_{\mathrm{piv}} + 1, \ldots, n \}. \tag{10}$$

Strategies (2) and (3) may of course be combined with strategy (1). The fourth strategy avoids the use of the diagonal elements of $\bar{C}$ and $\bar{L}$:

(4) perform no pivoting and estimate the smallest eigenvalue of A by *inverse iteration* applied implicitly to A:

$$v_0 = [0\ 0 \cdots 0\ 1]^{\mathrm{T}} \quad \text{or another initial guess} \tag{11a}$$

$$\left.\begin{aligned} \omega_i = A^{-1} v_{i-1} &= \begin{cases} (\bar{C}^{\mathrm{T}})^{-1} \bar{C}^{-1} v_{i-1}, \\ (\bar{L}^{\mathrm{T}})^{-1} \bar{D}^{-1} \bar{L}^{-1} v_{i-1}, \end{cases} \\ \lambda &= 1 / \| \omega_i \|_2, \\ v_i &= \lambda \omega_i \end{aligned}\right\} \quad i = 1, 2, \ldots \tag{11b}$$

where the computed factorizations are used in (11b).

Notice that the LINPACK condition estimator [2,3] can be considered as one specialized step of such inverse iteration. If this condition estimator is used to produce the initial guess in (11a) then one or two steps of the inverse iteration (11b) is often sufficient. The computational overhead associated with this strategy is therefore pn^2 flops, where p is a small integer, usually $p = 3$ or 4.

To experiment with the above strategies, a number of test matrices were generated by means of an eigenvalue decomposition of A:

$$A = Q\Lambda Q^{\mathrm{T}}. \tag{12}$$

Here, $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix consisting of the eigenvalues of A, and $Q \in \mathbb{R}^{n\times n}$ is a random orthogonal matrix. Thus, the condition number λ_1/λ_n of A as well as its type of rank deficiency can be specified explicitly via specification of Λ. Eigenvalue spectra were generated according to either an equidistant distribution of eigenvalues between $\lambda_1 = 1$ and $\lambda_n = 0$:

$$\lambda_i = (n-i)/(n-1), \quad i = 1, 2, \ldots, n \tag{13}$$

or a geometric distribution with eigenvalues decaying gradually from $\lambda_1 = 1$ towards $\lambda_n = 10^{-7}$:

$$\lambda_i = [10^{-7}]^{(i-1)/(n-1)}, \quad i = 1, 2, \ldots, n. \tag{14}$$

Distributions (13) and (14) give rise to matrices with well-determined and ill-determined numerical rank [6], respectively, and in both cases the numerical rank with respect to ϵ_{M} (i.e., the number of eigenvalues larger than $\epsilon_{\mathrm{M}}\lambda_1$) is $n-1$.

4. Experimental results

The experiments were carried out on a computer using chopped arithmetic and a single precision accuracy $\epsilon_{\mathrm{M}} = 1.19 \cdot 10^{-7}$. In all experiments, the order of the test matrices was $n = 64$, and the experiments were carried out in batches of 500 matrices (i.e., 500 matrices with the same Λ but random Q).

In each Cholesky and LDL^{T} factorization we computed all the quantities:

$$\rho^{(1)} = \min_k \{ c_{kk}^2/a_{kk} \}, \quad k = 1, 2, \ldots, n, \tag{15a}$$

$$\rho^{(2)} = \min_k \{ d_{kk}/a_{kk} \}, \quad k = 1, 2, \ldots, n, \tag{15b}$$

where c_{kk}^2 denotes the quantity in (2) immediately before taking the square root, and a_{kk} is the corresponding original diagonal element of A (if complete symmetric pivoting is applied to the nearly singular matrices in our experiments then $\rho^{(1)}$ and $\rho^{(2)}$ should be small positive or negative multiples of the machine precision ϵ_{M}). In order to evaluate the reliability of the above four strategies, each batch of 500 matrices was then factorized using these strategies, and in each case the largest value $\rho_{\max}^{(1)}$ and $\rho_{\max}^{(2)}$ were found. These quantities are experimental measures of the reliability of the particular strategy (the smaller number, the better).

The results from experiments with matrices with well-determined numerical rank (13) are shown in Table 1. As expected, it can be seen that complete symmetric pivoting is always a very reliable strategy. It can also be seen that a factorization without pivoting does not ensure the computation of a small diagonal element: for about 5% of the 500 test matrices, $\rho_{\max}^{(1)}$ and $\rho_{\max}^{(2)}$ were only of magnitude $\epsilon_{\mathrm{M}}^{1/2}$. Accumulation in double precision is seen to give slightly better

Table 1
The quantities $\rho^{(1)}$ and $\rho^{(2)}$ for test matrices with well-determined numerical rank

	Accumulation in single precision		Accumulation in double precision	
	$\rho^{(1)}_{\max}$	$\rho^{(2)}_{\max}$	$\rho^{(1)}_{\max}$	$\rho^{(2)}_{\max}$
Complete pivoting	$2.32 \cdot 10^{-6}$	$2.07 \cdot 10^{-6}$	$7.27 \cdot 10^{-7}$	$4.61 \cdot 10^{-7}$
No pivoting	$3.65 \cdot 10^{-3}$	$3.26 \cdot 10^{-3}$	$1.96 \cdot 10^{-3}$	$3.26 \cdot 10^{-6}$
Threshold pivoting with $\tau = 0.1$	$7.73 \cdot 10^{-5}$	$9.27 \cdot 10^{-5}$	$2.37 \cdot 10^{-5}$	$2.38 \cdot 10^{-5}$
Threshold pivoting with $\tau = 0.01$	$6.95 \cdot 10^{-4}$	$1.54 \cdot 10^{-3}$	$1.96 \cdot 10^{-3}$	$2.41 \cdot 10^{-4}$
Pivoting in the final $n_{\text{piv}} = 4$ steps	$1.12 \cdot 10^{-5}$	$2.09 \cdot 10^{-5}$	$3.61 \cdot 10^{-6}$	$2.11 \cdot 10^{-6}$

results, although it does not ensure the computation of small $\rho^{(1)}$ and $\rho^{(2)}$ either. Strategy (1) must therefore be rejected.

The second strategy with threshold pivoting requires that the threshold τ be surprisingly large in order to make the method reliable, and one can therefore not expect to reduce the number of pivoting considerably. Accumulation in double precision does not improve the reliability of this method very much, and strategy (2) must therefore also be rejected.

In the third strategy, we allowed pivoting in the final $n_{\text{piv}} = 4$ steps. This approach turned out to be astonishingly reliable for the present type of test matrices in which the eigenvector corresponding to the smallest eigenvalue λ_n is non-sparse, which means that the last rows of A are almost linear combinations of all the other rows of A. The technique is obviously not reliable in general, for example if the near-singularity of A is caused by a few of the first rows of A being almost identical. But the approach can for example be used in a 'hybrid' sparse factorization that switches to a non-sparse technique in the final steps of the factorization [9].

The experiments with strategy (4), using inverse iteration to estimate the smallest eigenvalue λ_n, gave much better results. Although rounding errors during the factorization, as we have seen, may hinder the computation of small diagonal elements in the computed $\bar{C}$, $\bar{L}$, and $\bar{D}$, these matrices *maintain* information about the near-singularity of A. To see this, we remind that the eigenvalues of the computed iteration matrices $\bar{C}\bar{C}^{\mathrm{T}}$ and $\bar{L}\bar{D}\bar{L}^{\mathrm{T}}$ in (11b) differ only slightly from those of A due to the following perturbation results from [5, Corollary 8.3-2]:

$$\left.\begin{aligned} |\lambda_i(\bar{C}\bar{C}^{\mathrm{T}}) - \lambda_i(A)| &\leqslant \| E^{(1)} \|_2 \\ |\lambda_i(\bar{L}\bar{D}\bar{L}^{\mathrm{T}}) - \lambda_i(A)| &\leqslant \| E^{(2)} \|_2 \end{aligned}\right\} \leqslant 2/3 n^{3/2} \epsilon_M \lambda_1 \approx 4.1 \cdot 10^{-5} \tag{16}$$

where we have used (6) and $\lambda_1 = 1$. Thus, although $\bar{C}$ and $\bar{D}$ may not have any small diagonal elements, the matrices $\bar{C}$, $\bar{L}$, and $\bar{D}$ are still guaranteed to reveal the near-singularity of A when used in the inverse iteration scheme (11b) to estimate λ_n. In our experiments, the inverse iterations always converged to an estimate $\bar{\lambda}_n$ of λ_n of magnitude

$$\bar{\lambda}_n \approx 10^{-7} \quad \text{to} \quad \bar{\lambda}_n \approx 10^{-8} \tag{17}$$

which is much better than could be expected from the bound (16), obtained from the conservative error analysis. Hence, the theory as well as our experimental results suggest inverse iteration as a very reliable and efficient technique for detection of a nearly singular matrix A. Also, notice

Table 2
The quantities $\rho^{(1)}$ and $\rho^{(2)}$ for test matrices with ill-determined numerical rank

	Accumulation in single precision		Accumulation in double precision	
	$\rho^{(1)}_{\max}$	$\rho^{(2)}_{\max}$	$\rho^{(1)}_{\max}$	$\rho^{(2)}_{\max}$
No pivoting	$3.64 \cdot 10^{-6}$	$3.67 \cdot 10^{-6}$	$3.70 \cdot 10^{-7}$	$3.69 \cdot 10^{-7}$

that if $\lambda_1 = \| A \|_2$ is estimated via an estimate of the norm of A (which can be done in $\frac{1}{2}n^2$ flops) then this technique yields as a by-product an estimate of the condition number λ_1/λ_n of A.

The similar experiments with matrices with ill-determined numerical rank (14) were not so extensive because it turned out that, even without pivoting, sufficiently small $\rho^{(1)}$ and $\rho^{(2)}$ were always found, cf. Table 2. This is due to the fact that the diagonal elements of $\bar{C}$ and $\bar{D}$—even without pivoting—tend to appear in decreasing order such that (7) is automatically satisfied.

We stress that determination of the numerical rank of A was not considered in these investigations—only the *detection* of a nearly singular matrix was considered. If the numerical rank is also to be determined in a reliable and efficient way, then an iterative technique as described Foster [4] and Chan [1] can be used.

5. Conclusion

The conclusions to be drawn form this experimental investigation of numerical methods for detection of near-singularity are:

(1) One can not expect a reliable detection of near-singularity simply by checking for small computed diagonal elements in a usual Cholesky or LDL^T factorization without symmetric pivoting.

(2) If the factorizations are performed with pivoting in the last few steps, then the check for small computed diagonal elements becomes more reliable although it is not 100% safe.

(3) In absence of pivoting, an accurate estimate of the smallest eigenvalue λ_n is guaranteed to be obtained from the computed factorization using inverse iteration (possibly combined with the LINPACK condition estimator), thus ensuring the detection of a nearly singular matrix.

Since the largest eigenvalue $\lambda_1 = \| A \|_2$ can always be estimated by a computationally cheap estimate of the norm of A, a by-product of this approach is the condition number of A which is useful also when A is non-singular. This technique requires only a small amount of computational overhead, usually about $3n^2$ or $4n^2$ flops, and it is therefore highly recommended for implementation in connection with sparse Cholesky and LDL^T factorizations.

References

[1] T.F. Chan, Rank revealing QR-factorizations, *Lin. Alg. Appl.*, to appear.

[2] A.K. Cline, C.B. Moler, G.W. Stewart and J.H. Wilkinson, An estimate for the condition number of a matrix, *SIAM J. Numer. Anal.* **16** (1979) 368–375.

[3] J.J. Dongarra, J.R. Bunch, C.B. Moler and G.W. Stewart, *LINPACK Users Guide* (SIAM Publications, Philadelphia, 1979).

[4] L.V. Foster, Rank and Null Space Calculations Using Matrix Decomposition without Column Interchanges, *Lin. Alg. Appl.* **74** (1986) 47–71.
[5] G.H. Golub and C.F. Van Loan, *Matrix Computations* (North Oxford Academic, 1983).
[6] P.C. Hansen, The truncated SVD as a method for regularization, *BIT*, to appear.
[7] R.S. Martin, G. Peters and J.H. Wilkinson, Symmetric decomposition of a positive definite matrix, in: J.H. Wilkinson and C. Reinsch, Eds., *Handbook for Automatic Computation. Vol. II, Linear Algebra* (Springer, Berlin, 1971).
[8] J. Meinguet, Refined error analysis of Cholesky factorization, *SIAM J. Numer. Anal.* **20** (1983) 1243–1250.
[9] N. Munksgaard, New factorization codes for sparse, symmetric and positive definite matrices, *BIT* **19** (1979) 43–52.