Bol. Soc. Esp. Mat. Apl. n^{o}0 (0000), 1–6

High relative accuracy algorithms for the symmetric eigenproblem

Juan Manuel Molera Molera *

Departamento de Matemáticas, Universidad Carlos III de Madrid

molera@math.uc3m.es

Abstract

In this talk we will review the basic facts and results in the field of high relative accuracy. We will see which algorithms and for which classes of matrices give high relative accuracy. In particular we present an algorithm [Dopico, Molera & Moro, SIAM J. Matrix Anal. Appl., 2003] that delivers high relative accuracy for the largest class of symmetric, definite and indefinite, matrices known so far. The algorithm is divided in two stages: the first one computes a SVD with high relative accuracy, and the second one obtains eigenvalues and eigenvectors from singular values and vectors. Using the SVD as a first stage is responsible both for the wide applicability of the algorithm and for being able to use only orthogonal transformations.

Palabras clave : *Symmetric eigenproblem, singular value decomposition, high relative accuracy.*

Clasificación por materias AMS : *65F15, 65G50, 15A18*

1. Introduction

An *orthogonal spectral decomposition* of a real symmetric n by n matrix A is a factorization $A = Q\,\Lambda\,Q^T$, where $Q = [q_1, \ldots, q_n]$ is real orthogonal and $\Lambda = \mathrm{diag}[\lambda_1 \geq \ldots \geq \lambda_n]$ is diagonal. Eigenvalues, $\widehat{\lambda}_i$, and eigenvectors, $\widehat{q}_i$, computed numerically with conventional methods, like QR, divide-and-conquer or bisection with inverse iteration have high *absolute* accuracy, i.e, they satisfy

$$|\lambda_i - \widehat{\lambda}_i| = O(\epsilon)\,\max_j |\lambda_j|, \tag{1}$$

for the eigenvalues, and

$$\Theta(q_i, \widehat{q}_i) = \frac{O(\epsilon)}{\frac{\min_{j\neq i}|\lambda_i-\lambda_j|}{\max_j|\lambda_j|}}, \tag{2}$$

*Grant BFM2003-00222

for the eigenvectors, being ϵ the unit roundoff of the finite arithmetic employed and $\Theta(q_i, \widehat{q}_i)$ the acute angle between vectors q_i and $\widehat{q}_i$. Thus, a conventional algorithm in general only guarantees correct digits in the eigenvalues with large enough magnitudes but may provide approximations for the small eigenvalues with no correct significant digits, or even with the wrong sign. Moreover, if there are two or more small eigenvalues, their eigenvectors may be computed very inaccurately, even when the eigenvalues are well-separated in the relative sense (e.g. $\lambda_i = 10^{-15}$ and $\lambda_j = 10^{-16}$ if $\lambda_1 = 1$).

High relative accuracy algorithms guarantee that computed eigenvalues have some correct digits, even if the eigenvalues have widely varying magnitudes. It is of interest to compute the tiniest eigenvalues with several correct digits because in some cases those are the ones that have physical meaning. The goal is to compute eigenvalues and eigenvectors with errors

$$|\lambda_i - \widehat{\lambda}_i| = O(\epsilon)|\lambda_i|, \tag{3}$$

instead of (1), and

$$\Theta(q_i, \widehat{q}_i) = \frac{O(\epsilon)}{relgap(\lambda_i)}, \tag{4}$$

instead of (2)[1]. The quantity $relgap(\lambda_i) = \min\left\{\min_{j\neq i} \frac{|\lambda_j - \lambda_i|}{|\lambda_i|}, 1\right\}$ is the eigenvalue relative gap.

High relative accuracy has been a very active area of research in the last fifteen years (see [3] and references therein for an overview). It is known that in general it is not possible to get (3) and (4) for an arbitrary matrix. At present, high relative accuracy can only be reached for certain classes of *symmetric* matrices. These are identified through a twofold approach. First it is necessary to determine classes of matrices and perturbations that allow high relative accuracy. In particular, a lot of work has been done in multiplicative perturbation theory (see [7] and references therein for a full review) and it is known that algorithms with small multiplicative backward errors will produce HRA. Second, algorithms that can use and exploit this special perturbation theory are to be supplied. Some of them will be review in this work.

It happens that the quest for high relative accuracy in the symmetric eigenproblem it is closely related to the high relative accuracy problem in the computation of singular values and vectors. Then we will devote section 2 to the SVD, and section 3 to the symmetric eigenvalue problem.

2. High Relative Accuracy and the Singular Value Decomposition

For a symmetric matrix the eigenvalue and the SVD problems are very closely related. HRA was first proven in 1990 for the SVD of bidiagonal

[1] For the sake of brevity in the rest of this work I will only speak of eigenvalues and singular values. There are equivalent results for eigenvectors and singular vectors.

matrices by Demmel and Kahan, by using zero a modified QR-type algorithm. Later on, Fernando and Parlett introduced a new algorithm (dqds, it is called) that also delivers HRA for the SVD problem. After that HRA has been proved for special classes: (Scaled) Cauchy, (Polynomial) Vandermonde, (Diag. Dominant) M-matrices, TN, ... (see [3] for a complete list of references). In 1999 Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač [3] proposed a unified algorithmic and theoretical approach for many High Relative Accuracy SVD computations. It was stablished that the necessary and suficient condition for a matrix to have a High Relative Accuracy SVD was that it had a Rank Revealing Decomposition (RRD) with small forward errors (see Theorem 1 below). Given $A \in \mathbb{C}^{m\times n}$, $m \geq n$, of rank r, $A = XDY^*$ is a RRD if $D \in \mathbb{C}^{r\times r}$ is diagonal and nonsingular, and $X \in \mathbb{C}^{m\times r}$ and $Y \in \mathbb{C}^{n\times r}$ are well conditioned matrices of full column rank[2]. The important thing it is that this condition can be checked using a finite computation (as GECP). Furthermore, they provided algorithms to implement it and were able to classify many classes of matrices allowing HRA SVD. The proposed algorithm has two steps

`1. Compute an RRD` $A = XDY^T$.

`2. Compute SVD of the RRD` $XDY^T = U\Sigma V^T$ `(using a Jacobi-type algorithm).`

Their main result establishes that for any matrices such that an accurate enough RRD can be obtained HRA SVD will be possible.

Teorema 1 *If the RRD factors, $A = XDY^* \approx \widehat{X}\widehat{D}\widehat{Y}^*$, are computed with the forward errors:*

$$|D_{ii} - \widehat{D}_{ii}| = O(\epsilon)|D_{ii}|, \quad \|X - \widehat{X}\| = O(\epsilon)\|X\|, \quad \|Y - \widehat{Y}\| = O(\epsilon)\|Y\|,$$

where ϵ is the unit roundoff, then the above algorithm compute SVD with high relative accuracy, i.e.,

$$|\widehat{\sigma}_i - \sigma_i| \leq O(\max\{\kappa(X), \kappa(Y)\}\epsilon)|\sigma_i|$$

3. HRA Algorithms for the symmetric eigenproblem

In terms of results and difficulty the high relative accuracy problem for the symmetric eigenvalue problem it is divided naturally in two cases: definite and undefinite matrices.

3.1. Positive definite matrices

The positive definite eigenproblem was the first to be treated from the point of view of getting HRA by a factorization+Jacobi rotations [4]. For matrices with good scaling properties HRA was proven by a two step algorithm:

`1. Compute a Cholesky decomposition` $A = LL^T$

`2. Compute SVD of` L `using right-Jacobi:` $LG_1G_2\cdots G_p = U\Sigma$

[2] The reduced SVD is an example of RRD.

The eigenvalues are $\Lambda = \Sigma^2$ and $Q = U$ the eigenvectors. It was proved that it is possible to get HRA for the class of positive definite matrices that are well scaled by its diagonal.

Teorema 2
Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite, then the eigenvalues $\{\widehat{\lambda}_i\}_{i=1}^n$ computed by previous Algorithm satisfy $|\widehat{\lambda}_i - \lambda_i| \leq O(\epsilon)\kappa(B)|\lambda_i|$, with $B = S^{-1}AS^{-1}$ and $S = diag[A_{11}^{1/2}, \dots, A_{nn}^{1/2}]$.

3.2. Indefinite symmetric matrices

The indefinite case is more complicated than the positive definite one. The first eigenproblem for which HRA was proved to be achievable was for scaled diagonally dominant matrices in 1990 [1]. A special technique was used: bisection plus inverse iteration. However, for the general case a similar approach (RRD + SVD) as in the positive definite case is possible though there are some important differences.

At present there are two available algorithms that can obtain HRA for different classes of matrices:

- `J-orthogonal hyperbolic algorithm` by Veselić and Slapničar (1992)[9, 8]
- `SVD +signs algorithm` by Dopico, Molera and Moro (2003)[5]

They are similar in that both use the two step aproach: Factorization + SVD (using Jacobi-type rotations). Both have computational cost higher than usual algorithms, though still $O(n^3)$ if the factorization step is $O(n^3)$. However they differ in the approach to the problem of the signs, and though they deliver similar (high relative) accuracy, the range of applicability and the bounds they get for the errors are different.

3.2.1. J-orthogonal algorithm

The two main steps of the `J-orthogonal algorithm` are

1. `Compute` $A = GJG^T$
2. `Apply the implicit` J`-orthogonal Jacobi method to the pair` (G, J):

$$GH_1H_2\cdots H_p = G_{p+1} = U\Delta \quad \text{with} \quad H_k^T J H_k = J$$

The eigenvalues are $\Lambda = \Delta^2 J$ and the eigenvectors $U = G_{p+1}\Delta^{-1}$ with $G_{k+1} = G_k H_k,\ \ k = 1:p,\ \ G_1 = G$.

The first step is a variation of the symmetric RRD $A = XDX^T$ Bunch-Parlett factorization [2], and the second step uses hyperbolic (J-orthogonal) implicit SVD Jacobi on GJG^T.

There are some objections to this method. First symmetric RRDs are less understood and more difficult to deal with than non-symmetric RRDs (GECP). Second, hyperbolic rotations have more arithmetic and its numeric properties are worse than orthogonal rotations (they do not preserve norms). Besides the routines that implement them are not included in state-of-the art software (`LAPACK`, for example). The accuracy result of this algorithm is

Teorema 3 *Let A be symmetric with $\{\lambda_i\}_{i=1}^n$ its eigenvalues and a symmetric High Relative Accuracy RRD given $A = XDX^T = GJG^T$ given. Let $\{\widehat{\lambda}_i\}_{i=1}^n$ be the eigenvalues computed by Step 2 of the* `J-Orthogonal Algorithm`*. Then*

$$\frac{|\widehat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon) \left[\kappa(X) + \max_{k=1:p+1} \kappa\left(G_k S_k^{-1}\right) \right]$$

where $G_k S_k^{-1}$ has columns of norm-2 equal to one.

The class of matrices for which HRA is achiveable depend on the scaling properties of the polar factor of the matrix A and also in the growth of the condition numbers of the intermediate matrices G_k.

3.3. Algorithm SVD with signs: SSVD

The other algorithm that computes eigenvalues and eigenvectors with HRA is based in two ideas:

1. Use the work and results in HRA in the computation of the SVD.
2. The singular values of a symmetric matrix are the absolute values of the eigenvalues.

The main improvements of this algorithm are that it can be used in the broadest class of matrices (any matrix for which a HRA SVD exists) yet and the use of orthogonal rotations in the SVD step.

The two main steps of this algorithm are:

```
1.  Compute an SVD of A = UΣV^T.
2.  Compute e-values, e-vectors from SVD.
```

The first step can be thought as having two steps if the SVD is computed with the algorithm in section 2. In the second step the eigenvalues and eigenvectors are obtained from singular values and vectors. It is easy to show how this is done in the case where all singular values are different.

Let $A = A^T = U\Sigma V^T \in \mathbb{R}^{n\times n}$ with, $i = 1, \ldots, n$, $\Sigma = \text{diag}[\sigma_i]$, $U = [u_i]$, $V = [v_i]$ with $v_i^T u_j = 0$ for $i \neq j$. The sign of the eigenvalues is obtained through the relation $V^T AV = V^T U\Sigma = \text{diag}[(v_i^T u_i)\sigma_i]$ and the eigenvectors are just the right (or the left) singular vectors $Q = V$. The presence of clusters of singular values (numerically probable) can be handled in a similar way.

The accuracy of this algorithm is stated as follows

Teorema 4 *Let A be symmetric with $\{\lambda_i\}_{i=1}^n$ its eigenvalues and an RRD with small forward errors given $A = XDY^T$. Then the computed eigenvalues $\{\widehat{\lambda}_i\}_{i=1}^n$ by the* `SSVD algorithm` *satisfy*

$$|\widehat{\lambda}_i - \lambda_i| \leq O(\epsilon)\ \max(\kappa(X), \kappa(Y))|\lambda_i|.$$

One problem that can be present in some cases is the accuracy of the eigenvectors. In principle it is affected by the singular value relative gap, intrisic in the singular vector computation. However, in most cases in can be fixed.

4. Conclusions

A lot of progress has been made in the field of High Relative Accuracy, specially for the SVD but also for the eigenvalue problem. The positive definite eigenproblem is well understood. Some more work has to be done in the indefinite eigenproblem. Though it is possible that the new release of `LAPACK` will have an HRA eigensolver available, specially when a very fast HRA Jacobi routine it is going to be included [6].

References

[1] J. Barlow and J. Demmel, *Computing accurate eigensystems of scaled diagonally dominant matrices*, SIAM J. Numer. Anal., 27 (1990), pp. 762-791.

[2] J. R. Bunch and B. N. Parlett, *Direct methods for solving symmetric indefinite systems of linear equations*, SIAM J. Numer. Anal., 8 (1971), pp. 639-655.

[3] J. Demmel, M. Gu, S. Eisenstat, I. Slapničar, K. Veselić and Z. Drmač, *Computing the singular value decomposition with high relative accuracy*, Linear Algebra Appl. 299 (1999), pp. 21-80.

[4] J. Demmel and K. Veselić, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1204-1245.

[5] F. M. Dopico, J. M. Molera and J. Moro, *An orthogonal high relative accuracy algorithm for the symmetric eigenproblem*, SIAM J. Matrix Anal. Appl., 25 (2003) pp. 301-351.

[6] Z. Drmač, *A posteriori computation of the singular vectors in a preconditioned Jacobi SVD algorithm*, IMA J. Numer. Anal., 19 (1999), pp. 191-213.

[7] I. Ipsen, *Relative perturbation results for matrix eigenvalues and singular values*, Acta Numerica, (1998), pp. 151-201.

[8] I. Slapničar, *Accurate Symmetric Eigenreduction by a Jacobi Method*, PhD Thesis, Fachbereich Mathematik Fernuniversität, Gesamthochschule Hagen, Germany, 1992.

[9] K. Veselić and I. Slapničar, *Floating-point perturbations of Hermitian matrices*, Linear Algebra Appl., 195 (1993), pp. 81-116.