A Relative Perturbation Bound for Positive Definite Matrices

Roy Mathias* and Krešimir Veselić†

December 18, 2013

Abstract

We give a sharp estimate for the eigenvectors of a positive definite Hermitian matrix under a floating-point perturbation. The proof is elementary.

Recently there have been a number of papers on eigenvector perturbation bounds that involve a perturbation of the matrix which is small in some relative sense, including the typical rounding errors in matrix elements ([1], [2],[9],[7],[3],[4]...). Some of these have complicated proofs and all of them involve the notion of 'the relative gap' between the eigenvalues. i.e. a relative distance of the unperturbed eigenvalue to the rest of the spectrum. Several such relative gaps are in use. Anyway, in any such estimate it is only the nearest eigenvalue that matters, one does not care for distant eigenvalues and their influence. Our bounds control primarily the matrix of the angles between the perturbed and the unperturbed eigenvectors, standard bounds with relative gaps may be derived from them any time. In particular, in our bounds the distant eigenvalues naturally damp out the perturbation of the corresponding components of the eigenvector. Or bounds are asymptotically sharp i.e. for small perturbations they reach the first term of the perturbation theory. Our proof is simple (of all works cited above ([4]) is closest to our idea) - its only technical tool is taking the square root of a positive definite matrix. The simplicity of our proof may make it useful in a classroom.

Theorem 1 *Let $H = U\Lambda U^*$ and $\widetilde{H} = H + \delta H = \widetilde{U}\widetilde{\Lambda}\widetilde{U}^*$ be positive definite. Assume that U and $\widetilde{U}$ are unitary and that Λ and $\widetilde{\Lambda}$ are diagonal. Let $S = U^*\widetilde{U}$ and assume*

$$\eta = \|H^{-1/2}\delta H H^{-1/2}\| < 1. \tag{1}$$

Then for any j and for any set $\mathcal{S}$ not containing j we have

$$\left(\sum_{i\in\mathcal{S}} |s_{ij}|^2\right)^{1/2} \le \max_{i\in\mathcal{S}} \frac{\lambda_i^{1/2}\widetilde{\lambda}_j^{1/2}}{|\lambda_i - \widetilde{\lambda}_j|}\frac{\eta}{\sqrt{1-\eta}} \tag{2}$$

*Department of Mathematics, College of William and Mary, Williamsburg, VA 23187. mathias@cs.wm.edu. Research supported in part by a grant from the National Science Foundation and a Summer Research Grant from the College of William and Mary.

†Lehrgebiet Mathematische Physik, Fernuniversität Hagen Postf. 940 D-58084 Hagen. kresimir.veselic@fernuni-hagen.de

and, in particular,

$$|s_{ij}| \leq \frac{\lambda_i^{1/2}\widetilde{\lambda}_j^{1/2}}{|\lambda_i - \widetilde{\lambda}_j|}\frac{\eta}{\sqrt{1-\eta}} \tag{3}$$

Proof: Let $X = \Lambda^{-1/2}U^*\delta H U\Lambda^{-1/2}$. Then

$$\begin{aligned} \|X\| &= \|\Lambda^{-1/2}U^*\delta H U\Lambda^{-1/2}\| \\ &= \|U\Lambda^{-1/2}U^*\delta H U\Lambda^{-1/2}U^*\| \\ &= \|H^{-1/2}\delta H H^{-1/2}\| \\ &= \eta. \end{aligned}$$

Also

$$\begin{aligned} \widetilde{U}\widetilde{\Lambda}\widetilde{U}^* &= \widetilde{H} \\ &= H + \delta H \\ &= U\Lambda^{1/2}(I + \Lambda^{-1/2}U^*\delta H U\Lambda^{-1/2})\Lambda^{1/2}U^* \\ &= U\Lambda^{1/2}(I+X)\Lambda^{1/2}U^* \end{aligned}$$

That is

$$\widetilde{U}\widetilde{\Lambda}\widetilde{U}^* = U\Lambda^{1/2}(I+X)\Lambda^{1/2}U^* \tag{4}$$

Now multiplying on the left by U^* and on the right by $\widetilde{U}$ gives

$$S\widetilde{\Lambda} = \Lambda S + \Lambda^{1/2}X\Lambda^{1/2}S\,.$$

Taking the ij element of this identity we obtain

$$s_{ij} = \frac{\lambda_i^{1/2}}{\widetilde{\lambda}_j - \lambda_i}(X\Lambda^{1/2}S)_{ij} \tag{5}$$

Thus, we need a bound on $X\Lambda^{1/2}S$. (4) can be written as

$$I = (\widetilde{\Lambda}^{-1/2}S^*\Lambda^{1/2})(I+X)(\Lambda^{1/2}S\widetilde{\Lambda}^{-1/2})$$

and hence

$$\|\Lambda^{1/2}S\widetilde{\Lambda}^{-1/2}\| = \|(I+X)^{-1}\|^{-1/2} \leq (1-\eta)^{-1/2}$$

Let $Y_{.j}$ denote the j-th column of the matrix Y. Then

$$\begin{aligned} \|(X\Lambda^{1/2}S)_{.j}\| &= \|(X\Lambda^{1/2}S\widetilde{\Lambda}^{-1/2})\widetilde{\Lambda}^{1/2})_{.j}\| \\ &= \widetilde{\lambda}_j^{1/2}\|(X\Lambda^{1/2}S\widetilde{\Lambda}^{-1/2})_{.j}\| \\ &\leq \widetilde{\lambda}_j^{1/2}\|X\|\ \|(\Lambda^{1/2}S\widetilde{\Lambda}^{-1/2})\| \\ &\leq \widetilde{\lambda}_j^{1/2}\eta(1-\eta)^{-1/2} \end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{i\in\mathcal{S}} |s_{ij}|^2 &= \sum_{i\in\mathcal{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \widetilde{\lambda}_j}\right)^2 |(X\Lambda^{1/2}S)_{ij}|^2 \\
&\le \max_{i\in\mathcal{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \widetilde{\lambda}_j}\right)^2 \sum_{j\in\mathcal{S}} |(X\Lambda^{1/2}S)_{ij}|^2 \\
&\le \max_{i\in\mathcal{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \widetilde{\lambda}_j}\right)^2 \|(X\Lambda^{1/2}S)_{.j}\|^2 \\
&\le \max_{i\in\mathcal{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \widetilde{\lambda}_j}\right)^2 \widetilde{\lambda}_j \eta^2 (1-\eta)^{-1}.
\end{aligned}$$

Taking square roots yields (2) which then implies (3). □

The last inequality can be used to obtain a bound for the perturbation of the eigenvectors. We can obviously always choose $\tilde{U}$ such that $s_{ii} \ge 0$ for all j. Then we have

$$||\tilde{U}_{\cdot j} - U_{\cdot j}|| = \sqrt{2}\sqrt{1 - s_{11}}$$

$$= \sqrt{2}\sqrt{1 - \sqrt{1 - |s_{1i}|^2 - \cdots - |s_{1,i-1}|^2 - s_|1, i+1|^2 - \cdots - |s_{1n}|^2}}.$$

Then by (2)

$$\begin{aligned}
||\tilde{U}_{\cdot j} - U_{\cdot j}|| &\le & (6)\\
&\le \sqrt{2}\sqrt{1 - \sqrt{1 - \frac{\eta^2}{1-\eta}\max_{i(\neq j)} \frac{\lambda_i \tilde{\lambda}_j}{(\lambda_i - \tilde{\lambda}_j)^2}}} & (7)\\
&\le \frac{\eta\sqrt{2}}{\sqrt{1-\eta}} \max_{i(\neq j)} \frac{\lambda_i^{1/2}\tilde{\lambda}_j^{1/2}}{|\lambda_i - \tilde{\lambda}_j|} & (8)\\
& & (9)
\end{aligned}$$

Different choices of $\mathcal{S}$ would allow estimates for invariant subspaces.

The bounds (2,3) involve $\widetilde{\lambda}_j$ and λ_j in a symmetric way. One can write the bounds entirely in terms of the eigenvalues of H by using the fact ([2, 5])

$$\lambda_j(1-\eta) \le \widetilde{\lambda}_j \le \lambda_j(1+\eta), \tag{10}$$

Since the proof is also very simple, we repeat it here for convenience. Note that (1) is equivalent to $|x^*\delta Hx| \le \eta x^* Hx$ for any vector x. Thus,

$$(1-\eta)x^*Hx \le x^*(H+\delta H)x \le (1+\eta)x^*Hx$$

and (10) follows from the monotonicity property for the eigenvalues.[1] If η is small then $\widetilde{\lambda}_j \approx \lambda_j$.[2] In this case the right hand side in (3) reads asymptotically

$$\frac{\lambda_i^{1/2}\lambda_j^{1/2}}{|\lambda_i - \lambda_j|}\eta. \tag{11}$$

On the other hand, for a simple eigenvalue λ_i the perturbation theory gives

$$u_i^*\widetilde{u}_j = \frac{u_i^*\delta H u_j}{\lambda_i - \lambda_j} + \text{ higher terms}\,.$$

Taking here e.g. $\delta H = \eta_0 \lambda_i^{1/2}\lambda_j^{1/2}(u_i u_j^* + u_j u_i^*)$, $\eta_0 < 1$, we obtain

$$u_i^*\widetilde{u}_j = \frac{\lambda_i^{1/2}\lambda_j^{1/2}}{\lambda_i - \lambda_j}\eta_0 + \text{ higher terms}\,.$$

Since here $\eta = \eta_0$ we see that our bound is asymptotically sharp.

Since the space dimension n does not enter the main estimates (2,3) they will hold for corresponding perturbations of Hilbert space operators with compact resolvent as well.

Relative bounds in the computational practice are seldom found in the genuine form (1). Most common is the componentwise bound

$$|\delta H_{ij}| \le \varepsilon |H_{ij}| \tag{12}$$

where ε is e.g. the machine precision. It can be shown (see e.g. [9]) that (12) implies (1) with

$$\eta = \||A|\|\, \|A^{-1}\|\varepsilon \le \sqrt{n}\kappa(A)\varepsilon\,, \tag{13}$$

where $\|\cdot\|$ is the spectral norm and D is any diagonal matrix. One chooses D so as to make the condition number $\kappa(A)$ as small as possible. As was shown in [8] a nearly optimal (within a factor $\sqrt{n}$) choice is $D = \sqrt{\mathrm{diag}(H)}$. A proof that the growth factor $\sqrt{n}$ in (13) cannot be dispensed with, is provided in [6].

For the convenience of the reader we will derive the first inequality in (13):

$$|x^*\delta H x| \le |x|^T|\delta H||x| \le \varepsilon |x|^T|A|D|x| \le \varepsilon \||A|\| x^* D^2 x \le \varepsilon \||A|\|\|A^{-1}\| x^* H x\,.$$

This implies (1) with η from (13).

Although the same type of eigenvector estimates can be expected for indefinite Hermitian matrices as well (cf. e.g. [9]), we were not able to extend our simple way of proof to this case as yet.

---

[1] The monotonicity is, of course, a consequence of the minimax theorem. Most perturbation estimates for the eigenvalues of Hermitian matrices, however, follow more directly from the monotonicity which, in turn, is a fact much easier to keep in mind and memorize. This may be of relevance in teaching this matter.

[2] Of course, here we assume that Λ and $\tilde{\Lambda}$ are equally ordered.

References

[1] Barlow, J., Demmel, J., Computing accurate i eigensystems of scaled diagonally dominant matrices, SIAM J. Numer. Anal. **27** (1988) 762-791.

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

[3] S.C. Eisenstat and I. C. Ipsen, Relative perturbation techniques for singular value problems. SIAM J. Num. Anal., to appear.

[4] R.-C. Li, Relative perturbation theory, UCLA Berkeley preprint 1994.

[5] R. Mathias, Spectral perturbation bounds for graded positive definite matrices. SIAM J. Matrix Anal. Appl. submitted for publication.

[6] R. Mathias, A perturbation bound for eigenvectors and for the matrix square root. *in preparation*, 1995.

[7] I. Slapničar and K. Veselić, Perturbations of the eigenprojections of factorised Hermitian matrices, Linear Algebra Appl. **218** (1995) 273–280.

[8] A. van der Sluis, Condition numbers and equilibrations of matrices, Numer. Math. **195** (1993) 14–23.

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