Perturbations of the Eigenprojections of a Factorised Hermitian Matrix

Ivan Slapničar*

University of Split
Faculty of Electrical Engineering, Mechanical Engineering
and Naval Architecture
R. Boškovića b.b
58000 Split, Croatia
e-mail *slap@fesb.hr*

and

Krešimir Veselić

Lehrgebiet Mathematische Physik
Fernuniversität Hagen
Postfach 940
D-58084 Hagen, Germany
e-mail *kresimir.veselic@fernuni-hagen.de*

Submitted by Ludwig Elsner

ABSTRACT

We give the perturbation bounds for the eigenprojections of a Hermitian matrix $H = GJG^*$, where G has a full column rank and J is non-singular, under the perturbations of the factor G. Our bounds hold, for example, when G is given with elementwise relative error. Our bounds contain relative gaps between the eigenvalues and may, thus, be much less pessimistic than the standard norm estimates.

In this paper we give the perturbation bounds for the eigenprojections of a Hermitian matrix

$$H = GJG^* \ , \tag{1}$$

where G is a $n \times r$ matrix of the full column rank, and J is a non-singular

*This work was done during the author's stay at Fernuniversität Hagen.

LINEAR ALGEBRA AND ITS APPLICATIONS xxx:1–xx (1992) 1

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

Hermitian matrix, under the perturbations of the factor G. More precisely, the perturbed matrix H' is defined by

$$H' = (G + \delta G)J(G + \delta G)^* \equiv G'J(G')^* \ , \tag{2}$$

where

$$\|\delta G x\|_2 \leq \eta \|Gx\|_2 \ . \tag{3}$$

The most common J is of the form

$$J = \begin{bmatrix} I_m & 0 \\ 0 & -I_{r-m} \end{bmatrix}$$

in which case m, $r - m$ and $n - r$ is the number of the positive, negative and zero eigenvalues of H, respectively. Such J appears in the indefinite symmetric decomposition which is the first step of an accurate algorithm for the eigenreduction of real symmetric matrices [7, 5], or is used as a preconditioner for indefinite systems [3]. The perturbation of the type (3) occurs, for example, whenever G is given with a floating-point error in the sense

$$|\delta G_{ij}| \leq \varepsilon |G_{ij}| \qquad \text{for all } i, j.$$

Then, as shown in [2, 6], (3) holds with

$$\eta = \frac{\sqrt{n}\varepsilon}{\sigma_{min}(B)} \ ,$$

where $B = GD$ and D is a non-singular diagonal scaling. The most usual (and nearly optimal) choice is to take the diagonal elements of D as the Euclidian norms of the columns of G [2, 6].

In [6] we proved that (2) and (3) imply

$$(1 - \eta)^2 \leq \frac{\lambda'_k}{\lambda_k} \leq (1 + \eta)^2 \ , \tag{4}$$

where λ_k and λ'_k are equally ordered eigenvalues of H and H', respectively. Our present result supplies the eigenvector counterpart of (4) and was mentioned as an open problem in [6].[†] The proof of our result for the simpler case of non-singular H is contained in the first author's dissertation [5].

As in [1, 2, 6], our estimates contain a factor called "relative gap" between an eigenvalue and the rest of the spectrum of H. To simplify the notation, as well as the statement and the proof of our main result, we

[†] Our paper, although closely related to [6], does not need the latter as a prerequisite.

assume that λ is positive. Negative eigenvalues of H are considered as the positive eigenvalues of the matrix $-H$. By λ_L and λ_R we denote the left and the right neighbour of λ in the spectrum $\sigma(H)$ of H, respectively. The relative gap $rg(\lambda)$ is defined as

$$rg(\lambda) = \min\left\{1, \frac{\lambda_R - \lambda}{\lambda_R + \lambda}, \frac{\lambda - \lambda_L}{\lambda + \lambda_L}\right\} .$$

Here the terms containing λ_L, λ_R appear in the expression above, if λ_L, λ_R exist and are positive, respectively. In this way very close eigenvalues may have large relative gaps, if they are absolutely small and, thus, our perturbation estimates may be much less pessimistic than the usual norm estimates. Note that our definition of relative gap is similar but not identical with those from [1, 2, 6].

The spectral projection belonging to a (possibly multiple) eigenvalue λ of H is given by [4]

$$P = \frac{1}{2\pi i}\int_\Gamma (\lambda I - H)^{-1} d\lambda , \tag{5}$$

where Γ is a curve around λ which separates λ from the rest of the spectrum of H. The perturbed spectral projection P' is obtained by interchanging H with H' in (5), while the integration path Γ remains unchanged. This means that the perturbation is small enough so that the contour Γ does not intersect the spectrum of

$$H_\kappa = (G + \kappa\delta G)J(G + \kappa\delta G)^* , \qquad 0 \le \kappa \le 1 .$$

This assumption is, in fact, contained in the assumptions of our theorem below and it implies that both projections P' and P have the same trace, that is, their ranges have the same dimension.

The key technical device of our proof is the simple fact that H and its pseudoinverse have the same set of eigenprojections and that the perturbation of the pseudoinverse can be conveniently expressed under the condition (3). We now state our main result:

THEOREM 1. *Let λ be a positive (possibly multiple) eigenvalue of a non-singular Hermitian matrix $H = GJG^*$ from (1), and let P be the corresponding eigenprojection. Let P' be the corresponding spectral projection of the perturbed matrix H' from (2) and (3). Then*

$$\|P' - P\|_2 \le \frac{4\bar\eta}{rg(\lambda)} \cdot \frac{1}{1 - \dfrac{3\bar\eta}{rg(\lambda)}} , \tag{6}$$

where $\bar\eta = \eta(2+\eta)$, provided that the right hand side in (6) is positive.

Proof. Denoting by $\mathcal{R}$, $\mathcal{N}$ the range and the null-space, respectively, we obviously have

$$\mathcal{R}(H) = \mathcal{R}(G) \ , \qquad \mathcal{N}(H) = \mathcal{N}(G^*) \ .$$

The inequality (3) is equivalent to

$$\|\delta G(G^*G)^{-1}G^*y\|_2 \le \eta \|y\|_2$$

for all $y \in \mathcal{R}(G)$. Note that under the conditions of the theorem both G^*G and G'^*G' are positive definite. We can obviously extend the above inequality to

$$\|\delta G(G^*G)^{-1}G^*\|_2 \le \eta \ . \tag{7}$$

The orthogonal projection P_0 onto $\mathcal{R}(H)$ is given by

$$P_0 = G(G^*G)^{-1}G^* \ . \tag{8}$$

An analogous formula holds for the perturbed projection P_0' onto $\mathcal{R}(H')$.

Let H^+ be the pseudoinverse of H, given by

$$HH^+ = H^+H = P_0 \ , \qquad P_0H^+ = H^+P_0 = H^+ \ .$$

The spectral projection P belonging to the eigenvalue λ of H can be written as

$$P = \frac{1}{2\pi i}\int_\Gamma S_\mu d\mu \ , \qquad S_\mu = (\mu I - H^+)^{-1} \ , \tag{9}$$

where Γ is now a curve around $1/\lambda$ which separates $1/\lambda$ from the rest of the spectrum of H^+. An analogous formula (again with the same Γ) holds for the perturbed projection P'. The proof of (9) uses the spectral decomposition of H and is omitted. By the same way one can prove the formula

$$H^+ = G(G^*G)^{-1}J(G^*G)^{-1}G^* \ ,$$

which also reflects the fact that $(G^*G)^{-1}G^*$ is the pseudoinverse of G.

We now show that

$$P = \frac{1}{2\pi i}\int_\Gamma GT_\mu G^* d\mu \ , \qquad T_\mu = (\mu G^*G - J)^{-1} \ . \tag{10}$$

Indeed,

$$\begin{aligned} S_\mu P_0 &= [\mu I - G(G^*G)^{-1}J(G^*G)^{-1}G^*]^{-1}G(G^*G)^{-1}G^* \\ &= G[\mu I - (G^*G)^{-1}J(G^*G)^{-1}G^*G]^{-1}(G^*G)^{-1}G^* \\ &= G(\mu G^*G - J)^{-1}G^* = GT_\mu G^* \ . \end{aligned} \tag{11}$$

Using this, (9) and the obvious identity

$$PP_0 = P_0P = P \ ,$$

we obtain (10). The similar identities for P' are obtained analogously.

Now

$$P' - P = \frac{1}{2\pi i}\int_\Gamma (G'T'_\mu G'^* - GT_\mu G^*)d\mu \ , \tag{12}$$

where

$$\begin{aligned} G'T'_\mu G'^* - GT_\mu G^* &= G(T'_\mu - T_\mu)G^* + \Phi \\ \Phi &= \delta G T'_\mu G^* + GT'_\mu \delta G^* + \delta G T'_\mu \delta G^* \ . \end{aligned} \tag{13}$$

Further,

$$G(T'_\mu - T_\mu)G^* = GT_\mu(T_\mu^{-1} - (T'_\mu)^{-1})T'_\mu G^* = GT_\mu \mu \Psi T'_\mu G^* \ ,$$

where

$$\Psi = -\delta G^* G - G^*\delta G - \delta G^* \delta G \ .$$

Using (11) and (8), we have

$$G(T'_\mu - T_\mu)G^* = \mu S_\mu P_0 \Delta G T'_\mu G^* \ , \tag{14}$$

where (see also (7))

$$\begin{aligned} \Delta &= G(G^*G)^{-1}\Psi(G^*G)^{-1}G^* = -\Delta_1^* P_0 - P_0\Delta_1 - \Delta_1^*\Delta_1 \\ \Delta_1 &= \delta G(G^*G)^{-1}G^* \ , \qquad \|\Delta_1\|_2 \le \eta \\ \|\Delta\|_2 &\le \bar\eta = 2\eta + \eta^2 \ . \end{aligned} \tag{15}$$

From (10), (11) and (15), it follows

$$\begin{aligned} GT'_\mu G^* &= G[\mu(G^* + \delta G^*)(G + \delta G) - J]^{-1}G^* \\ &= G(T_\mu^{-1} - \mu\Psi)^{-1}G^* = G(I - \mu T_\mu \Psi)^{-1}T_\mu G^* \\ &= G[I - \mu(G^*G)^{-1}G^*GT_\mu G^*G(G^*G)^{-1}\Psi]^{-1}(G^*G)^{-1}G^*GT_\mu G^* \\ &= G(G^*G)^{-1}G^*(I - \mu GT_\mu G^*\Delta)^{-1}S_\mu P_0 \\ &= P_0(I - \mu S_\mu P_0 \Delta)^{-1}S_\mu P_0 \ . \end{aligned}$$

Thus,

$$\|GT'_\mu G^*\|_2 \le \frac{z}{1 - \bar\eta w} \ , \tag{16}$$

where

$$w = \max_{\mu\in\Gamma} \|\mu S_\mu P_0\|_2 \ , \qquad\qquad z = \max_{\mu\in\Gamma} \|S_\mu P_0\|_2 \ . \tag{17}$$

From (13), (15) and (16) we obtain

$$\begin{aligned} \|\Phi\|_2 &\leq \|\Delta_1 G T'_\mu G^*\|_2 + \|G T'_\mu G^* \Delta_1^*\|_2 + \|\Delta_1 G T'_\mu G^* \Delta_1^*\|_2 \\ &\leq \bar\eta \|G T'_\mu G^*\|_2 \leq \bar\eta \frac{z}{1-\bar\eta w} \ . \end{aligned} \tag{18}$$

Combining (13), (14), (15) and (16-18), we obtain

$$\begin{aligned} \|G' T'_\mu G'^* - G T_\mu G^*\|_2 &\leq \|G(T'_\mu - T_\mu)G^*\|_2 + \|\Phi\|_2 \\ &\leq \|\mu P_0 S_\mu \Delta G T'_\mu G^*\|_2 + \bar\eta \|G T'_\mu G^*\|_2 \\ &\leq \bar\eta \frac{1+w}{1-\bar\eta w} \ . \end{aligned} \tag{19}$$

Taking Γ as a circle of radius r around $1/\lambda$, (12) and (19) give

$$\|P' - P\|_2 \leq r z \bar\eta \frac{1+w}{1-\bar\eta w} \ , \tag{20}$$

so it remains to estimate z and w. We have

$$\begin{aligned} w &= \max_{\mu\in\Gamma} \|\mu S_\mu P_0\|_2 = \max_{\mu\in\Gamma} \max_{\substack{\nu\in\sigma(H^+) \\ \nu\neq 0}} \frac{|\mu|}{|\mu-\nu|} \ , \\ z &= \max_{\mu\in\Gamma} \|S_\mu P_0\|_2 = \max_{\mu\in\Gamma} \max_{\substack{\nu\in\sigma(H^+) \\ \nu\neq 0}} \frac{1}{|\mu-\nu|} \ . \end{aligned}$$

The non-vanishing eigenvalues of H^+ are the inverses of the non-vanishing eigenvalues of H. Note the remarkable fact that, due to the presence of the projection P_0, zero eigenvalues do not enter the above formulae for w and z. Since Γ is a circle, the maxima in the above relations are attained for μ's which lie on the real axis.

If λ_R exists, then we choose r as

$$r = \frac{1}{2} \min \left\{ \frac{1}{\lambda} - \frac{1}{\lambda_R}, \frac{1}{\lambda_L} - \frac{1}{\lambda} \right\} \ ,$$

and if λ_R does not exist, then we choose r as

$$r = \frac{1}{2} \min \left\{ \frac{1}{\lambda}, \frac{1}{\lambda_L} - \frac{1}{\lambda} \right\} \ .$$

It is easy to see that we always have $z = 1/r$. Since $\mu = 1/\lambda \pm r$, we have

$$w = \max\left\{\frac{1/\lambda - r}{1/\lambda - r - 1/\lambda_R}, \frac{1/\lambda + r}{r}, \frac{1/\lambda + r}{1/\lambda_L - 1/\lambda - r}\right\} .$$

Now if $r = (1/\lambda - 1/\lambda_R)/2$, then

$$w = 1 + \frac{2}{\dfrac{\lambda_R - \lambda}{\lambda_R}} \leq 1 + \frac{2}{rg(\lambda)} \leq \frac{3}{rg(\lambda)} ,$$

and (6) follows by inserting this and $z = 1/r$ into (20).

If $r = (1/\lambda_L - 1/\lambda)/2$, then

$$w = \frac{\lambda - \lambda_L}{\lambda + \lambda_L} \leq \frac{1}{rg(\lambda)} ,$$

and (6) follows by inserting this and $z = 1/r$ into (20).

Finally, if $r = 1/(2\lambda)$ (λ_R does not exist), then $w = 3$ and (6) follows by inserting this and $z = 1/r$ into (20).

Positivity of the right hand side of (6) justifies, in turn, our choice of the same Γ in the definitions of P and P' in (12) as follows: (4) implies that $1/\lambda_R$ can increase to at most $1/(\lambda_R(1-\eta)^2)$, $1/\lambda_L$ can decrease to at least $1/(\lambda_L(1+\eta)^2)$ and the eigenvalues of $(H')^+$ which correspond to $1/\lambda$ remain in the interval $[1/(\lambda(1+\eta)^2), 1/(\lambda(1-\eta)^2)]$. Positivity of the right hand side of (6) always implies $rg(\lambda) > 6\eta$. This, together with our choice of r, implies that Γ contains no points of the spectrum of $(H')^+$ and that the interior of Γ contains exactly those eigenvalues of $(H')^+$ which correspond to $1/\lambda$. ■

Remark 2. It is possible to prove theorem similar to Theorem 1 for a cluster of eigenvalues, as well. All eigenvalues of the cluster must be either positive or negative. The relative gap for the cluster is then defined using λ_L (λ_R) and the leftmost (rightmost) member of the cluster, respectively. The $r \cdot z$ term of (20) is then larger than 1, and smaller than the inverse of the relative gap of the cluster.

We conclude the paper by giving the perturbation bounds for the eigenvectors corresponding to simple eigenvalues. Suppose that the assumptions of Theorem 1 are fulfilled, and that λ and λ' are both simple and non-zero (that is, they have the same sign). Let v and $v' = v + \delta v$ be the corresponding unit eigenvectors, and let ϕ be the angle between them. Then $P = vv^*$, $P' = v'(v')^*$, and $P' - P$ is a matrix of rank 2 with the non-trivial eigenvalues, say, γ_1 and γ_2. Since $\mathrm{Tr}\,(P' - P) = 0$, we have $|\gamma_1| = |\gamma_2| \equiv \gamma$. Now

$$2\gamma^2 = \mathrm{Tr}\,[(P' - P)(P' - P)] = 2\sin^2\phi \ ,$$

so that

$$||P' - P||_2 = |\sin\phi| \ .$$

This finally implies

$$||\delta v||_2 = 2|\sin(\phi/2)| \leq \sqrt{2}||P' - P||_2 \ .$$

Combining the above relation with (6), we obtain the bound on $||\delta v||_2$. We expect this bound to compare favourably to the corresponding bounds from [1, 2] since it does not contain the factors $(n-1)$ or $(n-1)^{1/2}$, respectively.

REFERENCES

1 J. Barlow, J. Demmel, *Computing Accurate Eigensystems of Scaled Diagonally Dominant Matrices*, SIAM J. Numer. Anal., Vol. 27, No. 3, (762-791) 1990.

2 J. Demmel, K. Veselić, *Jacobi's method is more accurate than QR*, SIAM J. Matr. Anal. Appl., Vol. 13, No. 4, (1204-1245) 1992.

3 P. H. Gill, W. Murray, D. B Ponceleón, M. A. Saunders, *Preconditioners for Indefinite Systems Arising in Optimization*, SIAM J. Matr. Anal. Appl., Vol. 13, No. 1, (292-311) 1992.

4 T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1966.

5 I. Slapničar, *Accurate Symmetric Eigenreduction by a Jacobi Method*, Ph. D. Thesis, Fernuniversität, Hagen 1992.

6 K. Veselić, I. Slapničar, *Floating-point perturbations of Hermitian matrices*, Seminarberichte aus dem Fachbereich Mathematik und Informatik Nr. 41, Fernuniversität, Hagen 1991, submitted to Lin. Alg. Appl.

7 K. Veselić, *A Jacobi eigenreduction algorithm for definite matrix pairs*, to appear in Numer. Math.