SIAM J. MATRIX ANAL. APPL.
Vol. 27, No. 1, pp. 82–86

HOW TO DEDUCE A PROPER EIGENVALUE CLUSTER FROM A PROPER SINGULAR VALUE CLUSTER IN THE NONNORMAL CASE*

STEFANO SERRA-CAPIZZANO†, DANIELE BERTACCINI‡, AND GENE H. GOLUB§

Abstract. We consider a generic sequence of matrices (the nonnormal case is of interest) showing a proper cluster at zero in the sense of the singular values. By a direct use of the notion of majorizations, we show that the uniform spectral boundedness is sufficient for the proper clustering at zero of the eigenvalues: if the assumption of boundedness is removed, then we can construct sequences of matrices with a proper singular value clustering and having all the eigenvalues of an arbitrarily big modulus. Applications to the preconditioning theory are discussed.

Key words. proper and weak clustering, majorization, preconditioning

AMS subject classifications. 65F10, 15A18, 15A12

DOI. 10.1137/040608027

1. The result. Let $\mathbf{x}$ be a generic real vector with entries x_j, $j = 1, \ldots, s$. We define $\mathbf{x}^{\downarrow}$ with entries $x_j^{\downarrow}$, $j = 1, \ldots, s$, to be the nonincreasing ordering of $\mathbf{x}$, i.e.,

$$x_1^{\downarrow} \geq x_2^{\downarrow} \geq \cdots \geq x_s^{\downarrow}$$

with $\mathbf{x}^{\downarrow} = P\mathbf{x}$ and P a suitable s-by-s permutation matrix. A vector $\mathbf{a}$ with real entries a_j, $j = 1, \ldots, s$, is weakly majorized (see, e.g., [3]) by a vector $\mathbf{b}$ with real entries b_j, $j = 1, \ldots, s$, if

$$\sum_{j=1}^{k} \mathbf{a}^{\downarrow} \leq \sum_{j=1}^{k} \mathbf{b}^{\downarrow} \quad \forall k = 1, \ldots, s;$$

in that case, we write $\mathbf{a} \prec_w \mathbf{b}$. Weyl's majorant theorem (see, e.g., [3, Theorem II.3.6]) establishes a fundamental majorization relation between the eigenvalues and the singular values of a given matrix $A \in M_s(\mathbf{C})$ ($M_s(\mathbf{C})$ denotes the space of the complex s-by-s matrices).

THEOREM 1.1. *Let $A \in M_s(\mathbf{C})$, let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_s$ be the singular values of A, and let $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_s|$ with λ_j, $j = 1, \ldots, s$, being the eigenvalues of A. Take $\phi : \mathbf{R}_+ \to \mathbf{R}_+$ such that $\phi(\exp(t))$ is convex and monotone nondecreasing. Then*

$$\phi(|\lambda|) \prec_w \phi(\sigma)$$

with $\phi(|\lambda|)$, $\phi(\sigma)$ real vectors of size s and whose jth entry, $j = 1, \ldots, s$, is given by $\phi(|\lambda_j|)$ and $\phi(\sigma_j)$, respectively.

*Received by the editors May 11, 2004; accepted for publication (in revised form) November 1, 2004; published electronically June 22, 2005.
http://www.siam.org/journals/simax/27-1/60802.html

†Dipartimento di Fisica e Matematica, Università dell'Insubria, Via Valleggio 11, 22100 Como, Italy (stefano.serrac@uninsubria.it, serra@mail.dm.unipi.it). The work of this author was partially supported by MIUR grant 2002014121.

‡Dipartimento di Matematica, Università di Roma "La Sapienza," P.le A. Moro 2, 00185 Roma, Italy (bertaccini@mat.uniroma1.it). The work of this author was partially supported by MIUR grant 2002014121.

§Department of Computer Science, Stanford University, Gates 2B, CA 94305 (golub@stanford.edu).

The former result is now used in a context of asymptotic linear algebra for deducing eigenvalue clustering from the singular value clustering: we recall that a sequence A_n is *properly clustered* at $\alpha \in \mathbf{R}$ if for every $\epsilon > 0$ there exists a pure constant c_ϵ such that

$$\begin{aligned} \#\{\text{singular values of } A_n \notin (\alpha-\epsilon, \alpha+\epsilon)\} &\le c_\epsilon \quad \text{(singular values proper clustering)},\\ \#\{\text{eigenvalues of } A_n \notin (\alpha-\epsilon, \alpha+\epsilon)\} &\le c_\epsilon \quad \text{(eigenvalues proper clustering)}. \end{aligned}$$

THEOREM 1.2. *Let $A_n \in M_n(\mathbf{C})$ be a sequence of uniformly spectrally bounded matrices such that the singular values are properly clustered at zero. Then the eigenvalues of A_n are properly clustered as well.*

Proof. Let $\lambda_n = (\lambda_1^{(n)}, \ldots, \lambda_n^{(n)})^T$ and $\sigma_n = (\sigma_1^{(n)}, \ldots, \sigma_n^{(n)})^T$ be the (ordered) vectors of the eigenvalues and singular values of A_n, respectively, where $\sigma_1^{(n)} \ge \sigma_2^{(n)} \ge \cdots \ge \sigma_n^{(n)}$ and $|\lambda_1^{(n)}| \ge |\lambda_2^{(n)}| \ge \cdots \ge |\lambda_n^{(n)}|$. From the assumptions we know that there exists M a positive constant independent of n, and, for every $\epsilon > 0$, there exists $\bar{n} = \bar{n}_\epsilon$ such that

$$\|A_n\| = \sigma_1^{(n)} \le M \quad \forall n \quad \text{(uniform boundedness)}, \tag{1.1}$$

$$\sigma_{\bar{n}}^{(n)} < \epsilon \quad \forall n \ge \bar{n} \quad \text{(singular value clustering at zero)}. \tag{1.2}$$

Now we would like to prove the spectral clustering at zero of the eigenvalues. In the normal case this is a trivial result since $\sigma_j^{(n)} = |\lambda_j^{(n)}|$, $j = 1, \ldots, n$, while, in the general case, the essential tool is Weyl's majorant theorem: Theorem 1.1 with $\phi(z) = z$. In such a way, for every $k = 1, \ldots, n$, we have

$$\sum_{j=1}^{k} |\lambda_j^{(n)}| \le \sum_{j=1}^{k} \sigma_j^{(n)}. \tag{1.3}$$

By contradiction, we suppose that the eigenvalues of A_n are not properly clustered at zero. Therefore there exists a sequence of positive integers α_n monotonically going to infinity, and there exists $c > 0$ (independent of n) such that $|\lambda_{\alpha_n}^{(n)}| > c$ at least for a subsequence $n = n_q$ (where n_q is a strictly increasing sequence of integers). As a consequence, for every q large enough, we deduce

$$\sum_{j=1}^{\alpha_{n_q}} |\lambda_j^{(n_q)}| > c\alpha_{n_q}$$

and, due to (1.1) and (1.2), simultaneously we have

$$\sum_{j=1}^{\alpha_{n_q}} \sigma_j^{(n_q)} < M\bar{n} + (\alpha_{n_q} - \bar{n})\epsilon, \qquad \alpha_{n_q} > \bar{n}.$$

Thus, by putting together (1.3) with $k = \alpha_{n_q}$ and the latter two inequalities, we find $c\alpha_{n_q} < M\bar{n} + (\alpha_{n_q} - \bar{n})\epsilon$ for every q large enough. Finally, by dividing by α_{n_q} and by making the limit as q tends to infinity (α_{n_q} will go to infinity as well), we conclude that for every $\epsilon > 0$, there exists $\bar{n} = \bar{n}_\epsilon$ (independent of $n = n_q$ and therefore of q) such that $c \le \epsilon$. Since c is a pure positive constant and $\epsilon > 0$ is arbitrary, the choice of $\epsilon = c/2$ leads to $c \le c/2$ with $c > 0$, which is a contradiction, and the proof is complete. $\square$

2. Discussion and applications. The first observation concerns the essentiality of the assumptions of Theorem 1.2. If the singular values of A_n are not properly clustered at zero and the matrices A_n are definitely normal, then the eigenvalues cannot be properly clustered at zero since the singular values coincide with the absolute value of the eigenvalues. More interestingly, we will see that the hypothesis of the spectral boundedness is also essential. Let ϵ_n be a sequence of positive numbers and let B_n be the sequence of matrices given by

$$B_n = \begin{pmatrix} 0 & \epsilon_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \epsilon_{n-1} \\ 0 & \dots & & \dots & 0 \end{pmatrix}, \ 1 \geq \epsilon_1 \geq \epsilon_2 \geq \dots \geq \epsilon_k > 0 \ \forall k \geq 2, \ \lim_{k\to\infty} \epsilon_k = 0.$$

A simple check shows that this sequence is spectrally clustered to zero in the sense of the eigenvalues ($\lambda_j^{(n)} = 0$ for every $j = 1, \dots, n$) and in the sense of the singular (since $\sigma_1^{(n)} = \epsilon_1 \geq \sigma_2^{(n)} = \epsilon_2 \geq \dots \geq \sigma_{n-1}^{(n)} = \epsilon_{n-1} > \sigma_n^{(n)} = 0$). Now we consider a rank one perturbation of B_n having an exploding norm. More precisely we define A_n as

$$A_n = B_n + \psi_n \mathbf{e}_n \mathbf{e}_1^T,$$

where $\mathbf{e}_k$ is the kth vector of the canonical basis of $\mathbf{C}^n$ and $\psi_n > 0$ such that

$$\lim_{n\to\infty} \left[\psi_n \left[\prod_{k=1}^{n-1} \epsilon_k \right] \right]^{1/n} = \infty.$$

By the Cauchy interlace theorem (see, e.g., [4]) applied to the singular values, we deduce immediately that the singular values of A_n are also properly clustered to zero (because the singular values of B_n are properly clustered to zero and A_n is a one-rank modification of B_n; see [8]). However, a direct computation proves that the characteristic polynomial of A_n coincides with

$$p_{A_n}(\lambda) = (-\lambda)^n - \psi_n \left[\prod_{k=1}^{n-1} \epsilon_k \right]$$

and therefore the eigenvalues of A_n have all the same modulus, which is given by

$$\left[\psi_n \left[\prod_{k=1}^{n-1} \epsilon_k \right] \right]^{1/n}.$$

Since the latter tends to infinity as n tends to infinity, we have that the eigenvalues of B_n not only are not clustered at zero in the proper sense but are also clustered at infinity in the modulus. The responsibility for this very pathological behavior is the huge norm of the one-rank correction and the high nonnormality of B_n: the combination of these two ingredient gives (as is well known to numerical analysts) a great sensitivity to the eigenvalues, whose global distribution shows a dramatic change, i.e., from a proper clustering to zero for B_n to a proper clustering to infinity for

A_n (see [5, section 7] for an illustration of the numerical difficulties related to similar examples).

The next step is to show some consequences of Theorem 1.2, and a very natural and important application is the study of the convergence behavior of Krylov methods when nonnormal iteration matrices are involved. In [2] we considered the preconditioning of non-Hermitian (but positive definite) matrices coming from convection-diffusion equations. If B_n denotes the sequence of the discretized problems and P_n is the sequence of preconditioners (Hermitian and positive definite), setting $B_n = \mathrm{Re}(B_n) + \mathbf{i}\mathrm{Im}(B_n)$, $\mathbf{i}^2 = -1$, $\mathrm{Re}(B_n)$, $\mathrm{Im}(B_n)$ Hermitian matrices (the real and the imaginary part of B_n, i.e., $\mathrm{Re}(B_n) = (B_n + B_n^*)/2$, $\mathrm{Im}(B_n) = (B_n - B_n^*)/(2\mathbf{i})$), we have proved the following:

F1 $P_n^{-1}\mathrm{Re}(B_n)$ has eigenvalues properly clustered to 1 and lying in a strictly positive uniformly bounded interval.

F2 $P_n^{-1}\mathrm{Im}(B_n)$ has eigenvalues properly clustered to 0 and lying in a uniformly bounded interval.

From these two items we prove the proper clustering at 1 of the eigenvalues of $P_n^{-1}B_n$ by using Theorem 1.2.

PROPOSITION 2.1. *With the previously given notation, and taking into account* F1 *and* F2*, we deduce that both the singular values and the eigenvalues of $P_n^{-1}B_n$ are properly clustered at* 1*. Moreover, all the eigenvalues of $P_n^{-1}B_n$ belong to a uniformly bounded rectangle with positive real part well separated from zero.*

Proof. We first observe that the matrices $P_n^{-1/2}\mathrm{Re}(B_n)P_n^{-1/2}$, $P_n^{-1/2}\mathrm{Im}(B_n)P_n^{-1/2}$ are both Hermitian since P_n is Hermitian positive definite (so that by the Schur canonical decomposition, its square root exists and is Hermitian positive definite) and $\mathrm{Re}(B_n)$, $\mathrm{Im}(B_n)$ are Hermitian by construction. Moreover, $P_n^{-1}\mathrm{Re}(B_n)$ is similar to $P_n^{-1/2}\mathrm{Re}(B_n)P_n^{-1/2}$ and $P_n^{-1}\mathrm{Im}(B_n)$ is similar to $P_n^{-1/2}\mathrm{Im}(B_n)P_n^{-1/2}$: as a consequence both the statements contained in F1 and F2 are true for $P_n^{-1/2}\mathrm{Re}(B_n)P_n^{-1/2}$ and $P_n^{-1/2}\mathrm{Im}(B_n)P_n^{-1/2}$, respectively. Therefore we can deduce properties on the singular values since the involved matrices are Hermitian (and a fortiori normal); more precisely, both $P_n^{-1/2}\mathrm{Re}(B_n)P_n^{-1/2} - I_n$ and $P_n^{-1/2}\mathrm{Im}(B_n)P_n^{-1/2}$ are spectrally uniformly bounded and properly clustered to zero in the singular value sense. From this, by using the field of values notion (see [4]), and since (by F1) the eigenvalues of $P_n^{-1/2}\mathrm{Re}(B_n)P_n^{-1/2}$ belong to a strictly positive uniformly bounded interval, all the eigenvalues of $P_n^{-1}B_n$ lie in a uniformly bounded rectangle with positive real part well separated from zero. Moreover, by a direct SVD inspection, we see that the sequence

$$P_n^{-1/2}\mathrm{Re}(B_n)P_n^{-1/2} - I_n + \mathbf{i}P_n^{-1/2}\mathrm{Im}(B_n)P_n^{-1/2}$$

is spectrally uniformly bounded and properly clustered to zero in the singular value sense. As a consequence, all the assumptions of Theorem 1.2 are fulfilled with $A_n = P_n^{-1/2}\mathrm{Re}(B_n)P_n^{-1/2} - I_n + \mathbf{i}P_n^{-1/2}\mathrm{Im}(B_n)P_n^{-1/2}$, and thus the sequence A_n is also properly clustered to zero in the sense of the eigenvalues. Finally, we complete the proof by noticing that $A_n + I_n$ is properly clustered to 1 in the eigenvalue sense and

$$P_n^{-1}B_n = P_n^{-1/2}(A_n + I_n)P_n^{1/2} \quad \text{(a similarity transformation).} \qquad \square$$

The conclusions are in some sense contained in the former proposition: the discussed result could be the key to proving eigenvalue clustering when the preconditioner is Hermitian positive definite but the original problem is nonnormal. Some of these situations occur when dealing with partial differential equations and structured matrices (see, e.g., [2, 7]); in particular, concerning the aforementioned applications, we stress that there exist many tools for proving the singular value clustering in the nonnormal case [8, 6, 7], but not so many for dealing with the eigenvalues: in this direction we must mention [9], where deep and beautiful results are proven with regard to the notion of general clustering, i.e., up to $o(n)$ outliers; unfortunately, from the viewpoint of a fine convergence analysis of Krylov methods, we remark that often only the proper clustering studied in this note is of interest. Indeed, the weak clustering can be useful only when there is an estimate of the number of outliers which is described by a mildly growing function (for instance, a poly-logarithm of n). Furthermore, in both the cases (proper clustering or poly-logarithmic number of outlying eigenvalues) and for practical purposes, attention has to be paid to the multiplicative constants: as a matter of fact, the practical dependency on $\epsilon > 0$ is such that it usually appears in a denominator, and this represents a delicate point in a convergence analysis especially in the partial differential equations context where ϵ can be a function of the finesse parameter and therefore of n. Finally, we conclude by observing that future work should investigate the direction of providing specific tools in the case where the preconditioning sequence is constituted by nonnormal matrices (an attempt is contained in Theorem 4.3 of [1]).

REFERENCES

[1] D. Bertaccini, G. Golub, and S. Serra Capizzano, *Spectral Analysis and Superlinear Convergence of a Preconditioned Iterative Method for the Convection-Diffusion Equation*, reprint SCCM-04-11, Stanford University, Stanford, CA.

[2] D. Bertaccini, G. H. Golub, S. Serra Capizzano, and C. Tablino Possio, *Preconditioned HSS methods for the solution of non-Hermitian positive definite linear systems and applications to the discrete convection-diffusion equation*, Numer. Math., 99 (2005), pp. 441–484.

[3] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.

[4] G. Golub and C. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1983.

[5] G. H. Golub and J. H. Wilkinson, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Rev., 18 (1976), pp. 578–619.

[6] S. Serra Capizzano, *Spectral behavior of matrix sequences and discretized boundary value problems*, Linear Algebra Appl., 337 (2001), pp. 37–78.

[7] S. Serra Capizzano, *Generalized locally Toeplitz sequences: Spectral analysis and applications to discretized partial differential equations*, Linear Algebra Appl., 366 (2003), pp. 371–402.

[8] E. E. Tyrtyshnikov, *A unifying approach to some old and new theorems on distribution and clustering*, Linear Algebra Appl., 232 (1996), pp. 1–43.

[9] E. E. Tyrtyshnikov and N. L. Zamarashkin, *On the eigen and singular value clusters*, Calcolo, 33 (1996), pp. 71–78.