On computation of Hölder condition numbers

Elizabeth A. Kalinina
Saint-Petersburg State University
Saint-Petersburg, Russia
e-mail: ekalinina69@gmail.com

ABSTRACT

In the paper, an algebraic method to find the largest dimension of Jordan block for a given complex matrix is suggested. The polynomial whose roots are the eigenvalues corresponding to the largest Jordan blocks is constructed. Neither the knowledge of Jordan form of the matrix nor its characteristic polynomial is required. The results presented in the paper can be used for calculating Hölder condition number, which is the measure of the eigenvalues' variation under small perturbations of elements of the matrix.

Keywords

Hölder condition number, Kronecker product, eigenvalues and eigenvectors.

1. INTRODUCTION

In the paper, an algorithm to find the dimension of the largest Jordan block of a square complex matrix A is considered. Also a method to construct a polynomial whose roots are the eigenvalues of A associated with the largest Jordan blocks is suggested. The proposed results can be applied for the computation of Hölder condition numbers of the eigenvalues.

2. PRELIMINARY RESULTS

For a given square complex matrix A (of dimension k) and a perturbation parameter ε, we write the perturbed matrix as $A+\varepsilon B$ for an arbitrary matrix B. In this case, it is known [1] that every eigenvalue and every eigenvector of $A+\varepsilon B$ admits an expansion in fractional powers of ε, where zero-th order term is the corresponding eigenvalue or eigenvector of unperturbed matrix A.

To analyze the eigenvalues of perturbed matrix Hölder condition number is used.

Definition. Hölder condition number for the eigenvalue λ of matrix A is defined by

$$\operatorname{cond}(\lambda) = (n_{\max}, \alpha),$$

where $n_{\max}$ is the dimension of the largest Jordan block associated with λ, and

$$\alpha = \max_{||B|| \le 1} \operatorname{spr}(\mathcal{Y} B \mathcal{X}).$$

Here spr denotes spectral radius, and the columns of matrix $\mathcal{X}$ (rows of matrix $\mathcal{Y}$) are linearly independent right (left) eigenvectors associated with λ, each corresponding to a Jordan chain of the largest length $n_{\max}$.

For all $c > 1$, the eigenvalues λ' of $A+\varepsilon B$ converging to λ as $\varepsilon \to +0$ satisfy the inequality [2]

$$|\lambda' - \lambda| < c\alpha^{1/n_{\max}} \varepsilon^{1/n_{\max}} \tag{1}$$

for all sufficiently small positive ε.

Denote

$$\mathcal{C}_A = E \otimes A - A \otimes E, \tag{2}$$

where $A \otimes B$ stands for Kronecker product of matrices A and B, and E is the unity matrix of size k.

Theorem 1. [5] The eigenvalues of matrix $\mathcal{C}_A$ are $\lambda_i - \lambda_j$, $(i, j = 1, 2, \ldots, k)$.

Hence, $\mathcal{C}_A$ always has zero eigenvalue.

Consider matrix equation $A\mathfrak{X} = \mathfrak{X}A^T$.

Let $X_{k^2 \times 1}$ be an eigenvector of matrix $\mathcal{C}$ corresponding to the zero eigenvalue. Let

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix}, \ X_j \in \mathbb{C}^k.$$

With vectors X_j $(j = 1, \ldots, k)$ we build up a matrix. This matrix

$$\mathfrak{X} = (X_1, X_2, \ldots, X_k)$$

satisfies the equation $A\mathfrak{X} = \mathfrak{X}A^T$. Conversely, if matrix $\mathfrak{X} = (X_1, X_2, \ldots, X_k)$ is a solution of equation $A\mathfrak{X} = \mathfrak{X}A^T$, then vector $X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix}$ is an eigenvector of matrix $\mathcal{C}$ corresponding to its zero eigenvalue.

To find Hölder condition number of eigenvalue λ it is necessary to know the structure of Jordan blocks corresponding to λ. But the computation of Jordan canonical form is difficult in itself, so it is important to calculate Hölder condition number without this construction. In this paper, it is suggested to compute $n_{\max}$ and to construct a polynomial whose roots are the eigenvalues of A associated with the largest Jordan blocks if only Jordan blocks corresponding to the zero eigenvalue of matrix $\mathcal{C}$ are known.

3. DESCRIPTION OF THE ALGORITHM

Here we present theorems that allow us to find the size of the largest Jordan block of the matrix and the eigenvalues corresponding to the largest Jordan blocks.

3.1 Order of the largest Jordan block

Denote by $\lambda_1, \ldots, \lambda_k$ the eigenvalues of matrix A (some of them can be equal).

If A_J denotes Jordan form of matrix A, then $A = RA_JR^{-1}$ and $X' = R^{-1}X(R^T)^{-1}$. Hence, equation $AX = XA^T$ is equivalent to the equation

$$A_JX' - X'A_J^T = \mathbb{O}_{k\times k},$$

where $\mathbb{O}_{k\times k}$ stands for zero matrix of size k. Thus, it is sufficient to consider the case when $A = A_J$, i.e., matrix A is in Jordan normal form.

To find the size of the largest Jordan block we need to consider matrix (2).

First, consider the case when matrix A consists of a single Jordan block of size k, i.e.,

$$A = J = \begin{pmatrix} \lambda & 0 & 0 & \ldots & 0 & 0 \\ 1 & \lambda & 0 & \ldots & 0 & 0 \\ 0 & 1 & \lambda & \ldots & 0 & 0 \\ \ldots \\ 0 & 0 & 0 & \ldots & 1 & \lambda \end{pmatrix}_{k\times k}.$$

Denote by I the following matrix

$$I = \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \ldots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}_{k\times k}.$$

Then we have $\mathcal{C}_J^n =$

$$= \begin{bmatrix} I^n & \mathbb{O} & \ldots \mathbb{O} \\ -\mathsf{C}_n^1 I^{n-1} & I^n & \ldots \mathbb{O} \\ \mathsf{C}_n^2 I^{n-2} & -\mathsf{C}_n^1 I^{n-1} & \ldots \mathbb{O} \\ \ldots \\ (-1)^{k-1}\mathsf{C}_n^{k-1} I^{n-k+1} & (-1)^{k-2}\mathsf{C}_n^{k-2} I^{n-k+2} & \ldots I^n \end{bmatrix}. \tag{3}$$

Here $\mathsf{C}_n^j = \dfrac{n!}{j!(n-j)!}$. (If $n < k$, there are zeros in the lower left corner of the matrix.)

Theorem 2. The system of linear equations

$$\mathcal{C}_J^n \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} = \mathbb{O}_{k^2\times 1} \tag{4}$$

has exactly

$$nk - \mathsf{C}_{n+1}^2 + (n - \lfloor n/2 \rfloor)(\lfloor n/2 \rfloor + 1) \tag{5}$$

linear independent solutions.

Proof. First, consider the case $k \geq n-1$. Denote by V_j $(j = 1, 2, \ldots, k)$ vectors of Jordan chain of matrix J. We have $V_{j-i} = I^iV_j$. Obviously, all vectors X_j are the root vectors. Then every vector X_j is a linear combination of vectors V_i $(i = 1, \ldots, k)$:

$$\begin{aligned} &X_1 = \alpha_{11}V_1 + \alpha_{12}V_2 + \ldots + \alpha_{1n}V_n, \\ &X_2 = \alpha_{21}V_1 + \alpha_{22}V_2 + \ldots + \alpha_{2n}V_n + \alpha_{2,n+1}V_{n+1}, \\ &\ldots, \\ &X_{k-n+1} = \alpha_{k-n+1,1}V_1 + \alpha_{k-n+1,2}V_2 + \ldots + \alpha_{k-n+1,k}V_k, \\ &\ldots, \\ &X_k = \alpha_{k1}V_1 + \alpha_{k2}V_2 + \ldots + \alpha_{kk}V_k. \end{aligned}$$

Now we can rewrite system (4):

$$\begin{cases} \alpha_{1n} - \alpha_{2,n+1}\mathsf{C}_n^{n-1} = 0, \\ \alpha_{1,n-1} - \alpha_{2n}\mathsf{C}_n^{n-1} + \alpha_{3,n+1}\mathsf{C}_n^{n-2} = 0, \\ \alpha_{1,n} - \alpha_{2,n+1}\mathsf{C}_n^{n-1} + \alpha_{3,n+2}\mathsf{C}_n^{n-2} = 0, \\ \ldots, \\ \alpha_{n+1,k-1} - \alpha_{n+2,k}\mathsf{C}_n^{n-1} = 0. \end{cases}$$

And the last system of equations can be divided into subsystems as follows.

Let $1 \leq i \leq n$. The i-th subsystem contains unknowns

$$\alpha_{1,n-i+1}, \alpha_{2,n-i+2}, \ldots, \alpha_{k-n+i,k}.$$

There are $k-i$ equations in this subsystem. If $n \geq 2i$, the i-th subsystem has only zero solution. Otherwise the subsystem has $2i-n$ linear independent solutions.

Let $n+1 \leq i \leq k$. The i-th subsystem contains unknowns

$$\alpha_{i-n+1,1}, \alpha_{i-n+2,2}, \ldots, \alpha_{k,k+n-i}.$$

There are $k-i$ equations in this subsystem. The subsystem has n linear independent solutions.

There are also unknowns

$$\alpha_{k-n+2,1}, \alpha_{k-n+3,1}, \alpha_{k-n+3,2}, \alpha_{k-n+4,1}, \alpha_{k-n+4,2},$$

$$\alpha_{k-n+4,3}, \ldots, \alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{k,n-1}.$$

These unknowns can possess arbitrary values. There are C_n^2 such unknowns.

By summing all numbers of linear independent solutions we get the following formula for the case $k \geq n$:

$$nk + \sum_{j=\lfloor n/2 \rfloor+1}^{n} (2j-n) - \mathsf{C}_{n+1}^2.$$

Rewriting this formula, we obtain the statement of the theorem.

When $n > k$ we have

$$\begin{aligned} &X_1 = \alpha_{11}V_1 + \alpha_{12}V_2 + \ldots + \alpha_{1k}V_k, \\ &\ldots, \\ &X_k = \alpha_{k1}V_1 + \alpha_{k2}V_2 + \ldots + \alpha_{kk}V_k. \end{aligned}$$

In perfect analogy with the considered case, we get that the number of solutions of the system equals

$$\sum_{\ell=\lceil \frac{2k-n}{2} \rceil}^{2k-n-1} (2\ell + n - 2k) + k(n-k+1) + n\frac{2k-n-1}{2}.$$

Taking into account that

$$\left\lceil \frac{2k-n}{2} \right\rceil = \left\lceil k - \frac{n}{2} \right\rceil = k - \left\lfloor \frac{n}{2} \right\rfloor,$$

we obtain the statement of the theorem. □

And this is the corollary to Theorem 2.

Theorem 3. *There are k Jordan blocks of matrix $\mathcal{C}_J$ corresponding to Jordan block J of matrix A. Their dimensions are*

$$1, 3, 5, 7, \ldots, 2k-1.$$

Now consider arbitrary matrix A_J, i.e., Jordan form of matrix A. In this case, every vector X_k consists of subvectors X_{kj} $(j = 1, 2, \ldots, p)$. The dimension of X_{kj} is equal to the dimension of Jordan block of matrix A_J that is the k-th counting from the left upper corner. And to find the root vectors we have kp subsystems of every system of linear equations. Every such subsystem contains the components of the only one vector X_{kj}. Every subsystem has form (3). Therefore, for every Jordan block of dimension q we obtain q Jordan blocks of dimensions $1, 3, 5, 7, \ldots, 2q-1$. But there can be the other Jordan blocks in the case, when there exist two or more Jordan blocks corresponding to one eigenvalue of matrix A. Hence, the following theorem holds.

Theorem 4. *The largest dimension of matrix A Jordan block is equal to $n_{\max} = (s+1)/2$, where s is the largest dimension of matrix $\mathcal{C}_A$ Jordan block corresponding to the zero eigenvalue.*

Proof. Let p be the dimension of the largest Jordan block. Considering the structure of matrix $\mathcal{C}_{A_J}$ we obtain that for its zero eigenvalue, matrix $\mathcal{C}_A$ does not have Jordan blocks of dimension larger than $2p-1$. □

It is clear now that matrix A is diagonalizable iff $n_{\max} = 1$.

3.2 Eigenvalues corresponding to the largest Jordan blocks

The following theorems allow us to construct a polynomial whose roots are equal to the eigenvalues of matrix A corresponding to the largest Jordan blocks.

Theorem 5. *[6] For every eigenvalue λ of matrix A, there exists matrix D such that* $\mathrm{rank} D = 1$ *and* $AD = DA^T$. *The columns and transposed rows of D are the eigenvectors of matrix A corresponding to λ. For every eigenvalue λ, there are exactly u^2 such matrices D, where u stands for the geometric multiplicity of λ.*

Suppose that matrix $\mathcal{C}$ has t eigenvectors corresponding to the zero eigenvalue that makes the largest Jordan chains. Denote by $\mathfrak{C}_1, \mathfrak{C}_2, \ldots, \mathfrak{C}_t$ these vectors. Let us construct matrix $\mathfrak{C} = (\mathfrak{C}_1, \mathfrak{C}_2, \ldots, \mathfrak{C}_t)$.

Theorem 6. *Eigenvalues of matrix A corresponding to the largest Jordan blocks of dimension $n_{\max}$ are the roots of the equation*

$$\det(\mathfrak{C}^T(A \otimes E)\mathfrak{C} - \lambda \mathfrak{C}^T\mathfrak{C}) = 0. \qquad (6)$$

Proof. Consider eigenvalue λ of matrix A corresponding to the largest Jordan block of size $n_{\max}$.

Since $AD = DA^T$ for matrix D whose columns and transposed rows are the eigenvectors of A, then D is a linear combination of t matrices constructed of the eigenvectors of A: $\mathfrak{C}_1, \mathfrak{C}_2, \ldots, \mathfrak{C}_t$. In other words, for eigenvalue λ there exists matrix D corresponding to vector

$$\alpha_1\mathfrak{C}_1 + \alpha_2\mathfrak{C}_2 + \ldots + \alpha_t\mathfrak{C}_t$$

such that $AD = DA^T$.

Denote by $\mathfrak{A}$ vector $(\alpha_1, \alpha_2, \ldots, \alpha_t)^T$.

Now rewrite the equation recording matrix D as a vector:

$$(A \otimes E)\mathfrak{C}\mathfrak{A} = \lambda\mathfrak{C}\mathfrak{A}.$$

Left multiplying the both sides of this equation by $\mathfrak{C}^T$, we get

$$(\mathfrak{C}^T(A \otimes E)\mathfrak{C} - \lambda\mathfrak{C}^T\mathfrak{C})\mathfrak{A} = \mathbb{O},$$

and the last equation has non-zero solution $\mathfrak{A}$. It is possible iff

$$\det(\mathfrak{C}^T(A \otimes E)\mathfrak{C} - \lambda\mathfrak{C}^T\mathfrak{C}) = 0,$$

i.e., λ is a root of equation (6).

The equation does not have any other roots, because $\mathrm{rank}\,\mathfrak{C} = t$. □

Remark. It is proved in [6] that

$$\det(\mathfrak{C}^T(A \otimes E)\mathfrak{C} - \lambda\mathfrak{C}^T\mathfrak{C}) = \det(\mathfrak{C}^T(E \otimes A)\mathfrak{C} - \lambda\mathfrak{C}^T\mathfrak{C}).$$

4. NUMERICAL EXAMPLE

Let us find the dimension of the largest Jordan block and eigenvalues corresponding to the largest Jordan blocks for the matrix

$$A = \begin{pmatrix} -1 & 3 & -1 & 1 & 0 & 0 \\ -3 & 5 & -1 & 0 & 1 & 0 \\ -3 & 3 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 & 2 & -3 \\ 0 & 0 & 0 & 4 & 10 & -12 \\ 0 & 0 & 0 & 3 & 6 & -7 \end{pmatrix}.$$

The size of matrix $\mathcal{C}_A$ is 36, and the only one eigenvector corresponding to the zero eigenvalue is

$$\mathfrak{C}_1^T = (49, 70, 63, 0, 0, 0, 70, 100, 90, 0, 0, 0, 63, 90, 81,$$

$$0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0).$$

From equation (6) we obtain the eigenvalue corresponding to the largest Jordan block (in our case its dimension is 3):

$$105\,800\lambda - 211\,600 = 0,$$

so $\lambda = 2$.

Check. Matrix A has eigenvalues $\lambda = 2$ of multiplicity 5 and $\lambda = 1$ of multiplicity 1. In Jordan form there are three Jordan blocks for $\lambda = 2$ (their sizes are 3, 1 and 1) and one Jordan block for eigenvalue $\lambda = 1$ (its size is 1).

5. CONCLUSION

In the paper, an algorithm to determine the size of the largest Jordan block for a given complex matrix is suggested. The method allows us to find the eigenvalues associated with the largest Jordan blocks too. The results can be used for the evaluation of variation for the eigenvalues of the matrix under small perturbations of matrix elements.

REFERENCES

[1] T. Katō, "Perturbation Theory for Linear Operators". Berlin: Springer, 1980. 619 p.

[2] J. Moro, J.V. Burke, M.L. Overton, "On the Lidskii-Vishik-Lyusternik Perturbation Theory for Eigenvalues of Matrices with Arbitrary Jordan Structure", *SIAM J. Matrix Anal. Appl.* Vol. 18, N 4. Pp. 793–817, 1997.

[3] V.B. Lidskii, "Perturbation theory of non-conjugate operators", *USSR Computational Mathematics and Mathematical Physics*, 6:1, pp. 73–85, 1966.

[4] F. Chatelin, "Eigenvalues of Matrices". New York: John Wiley, 1993. 458 p.

[5] C.C. MacDuffee, "The Theory of Matrices". New York: Chelsea Publishing Company, 1956. 110 p.

[6] E.A. Kalinina, "The common eigenvalues of two matrices", *Dalnevost. matem. zhurn.*, Vol.13, N 1. Pp. 52–60, 2013.