Chin. Ann. of Math.
4B (4) 1983

SENSITIVITY OF THE EIGENVALUES OF A DEFECTIVE MATRIX

WU JINGKUN (吴景琨)
(*Fudan University*)

Abstract

This paper is concerned with the sensitivity of the eigenvalues of a defective matrix under small perturbations. The given estimate generalizes all special results of Wilkinson, Stewart, Bauer and Fike, when the eigenvalue is simple and when the matrix is nondefective, and interpretes the phenomenon indicated by Golub and Wilkinson for Multiple eigenvalues

In [1], with a quasi-Jorden block as an example, Golub and Wilkinson indicated that it is possible for a matrix to be highly defective without its eigenvalues being unduly sensitive. However, it hasn't been interpreted yet.

This paper gives an estimate of sensitivity of the eigenvalues for defective matrix (so that for general matrix). Not only it generalizes all special results of Wilkinson, Stewart, Bauer and Fike, when the eigenvalue is simple and when the matrix is nondefective, but also it interprets this phenomenon.

1. Eigenvalue Matrix

We shall denote the set of all $m\times n$ complex matrices by $\mathbf{C}^{m\times n}$.

Definition. *A matrix $A_\lambda\in\mathbf{C}^{r\times r}$ is called the eigenvalue matrix if it has an eigenvalue λ of multiplicity r.*

The following properties can be deduced by the definition.

1° Let $A\in\mathbf{C}^{n\times n}$ have an eigenvalue λ of multiplicity r. Then

(i) there exist eigenvalue matrices A_λ, $B_\lambda\in\mathbf{C}^{r\times r}$ and column full rank matrices G, $F\in\mathbf{C}^{n\times r}$ such that

$$AG=GA_\lambda,\ F^HA=B_\lambda F^H,$$

(ii) if there are also eigenvalue matrices A'_λ, $B'_\lambda\in\mathbf{C}^{r\times r}$ and column full rank matrices G', $F'\in\mathbf{C}^{n\times r}$ such that

$$AG'=G'A'_\lambda,\ F'^HA=B'_\lambda F'^H,$$

then there must exist nonsingular matrices S, H, Y, $Z\in\mathbf{C}^{r\times r}$ that satisfy

$$A'_\lambda=Y^{-1}A_\lambda Y,\ B'_\lambda=Z^{-1}B_\lambda Z,$$
$$G'=GS,\ F'=FH.$$

Manuscipt received September 4, 1981. Revised October 26, 1981.

2° Let $\prod_{i=1}^{s}(\lambda-\lambda_i)^{r_i}$ be the characteristic polynomial of $A\in\mathbf{C}^{n\times n}$, where all λ_i are different. Suppose that there are the eigenvalue matrices A_{λ_i}, $B_{\lambda_i}\in\mathbf{C}^{r_i\times r_i}$ and column full rank matrices G_i, $F_i\in\mathbf{C}^{n\times r_i}$ such that

$$AG_i=G_iA_{\lambda_i},\ F_i^HA=B_{\lambda_i}F_i^H,\ i=1,\ 2,\ \cdots,\ s.$$

Then the matrices $(G_1|G_2|\cdots|G_s)$, $(F_1|F_2|\cdots|F_s)$ are nonsingular and equalities

$$A(G_1|G_2|\cdots|G_s)=(G_1|G_2|\cdots|G_s)\ \mathrm{diag}\ [A_{\lambda_1},\ A_{\lambda_2},\ \cdots,\ A_{\lambda_s}],$$
$$(F_1|F_2|\cdots|F_s)^HA=\mathrm{diag}[B_{\lambda_1},\ B_{\lambda_2},\ \cdots,\ B_{\lambda_s}]\ (F_1|F_2|\cdots|F_s)^H$$

hold.

3° Let $A\in\mathbf{C}^{n\times n}$ has an eigenvalue λ of multiplicity r. Then there are eigenvalue matrix $A_\lambda\in\mathbf{C}^{r\times r}$ and two column full rank matrices G, F such that

$$AG=GA_\lambda,\ F^HA=A_\lambda F^H,\ F^HG=I.$$

And for A_λ, G and F are of one-one correspondence.

From 1° and 2°, it follows that the eigenvalue matrix is a generalization of the simple eigenvalue and the corresponding column full rank matrix is a generalization of the eigenvector. Moreover, each eigenvalue λ of A of multiplicity r corresponds to two subspaces $N((A-\lambda I)^r)$ and $N((A^H-\bar{\lambda}I)^r)$ which are called *the right and the left eigenvalue subspaces of* λ respectively. The columns of G and F form their basis respectively. A_λ and B_λ are uniquely determined by G and F respectively.

We say that the columns of G and F form a *dual basis pair* of the eigenvalue subspaces in 3°. For A_λ, G and F are not unique, but A_λ is uniquely determined by G and F.

2. Perturbation of the eigenvalue subspace

Using the sufficient condition (see [2]) under which the matrix equation

$$XM-NX=K-XLX$$

has a solution, we can prove the following

Theorem 1. *Let $A\in\mathbf{C}^{n\times n}$. Let $\prod_{i=1}^{s}(\lambda-\lambda_i)^{r_i}$ be the characteristic polynomial of A, where all λ_i are different. Suppose that there are eigenvalue matrices $A_{\lambda_i}\in\mathbf{C}^{r_i\times r_i}$ and column full rank matrices G_i, $F_i\in\mathbf{C}^{n\times r_i}$ such that*

$$AG_i=G_iA_{\lambda_i},\ F_i^HA=A_{\lambda_i}F_i^H,\ F_i^HG_i=I,\ i=1,\ 2,\ \cdots,\ s.$$

And for $i=1,\ 2,\ \cdots,\ s$, set

$$k_i=\|F_i\|_F\cdot\|G_i\|_F,\ k_i'=\|\mathrm{complem}(F_i)\|_F\cdot\|\mathrm{complem}(G_i)\|_F,$$
$$s_i=\min_{j\neq i}\mathrm{sep}_F\ (A_{\lambda_i},\ A_{\lambda_j}),$$

where $\mathrm{complem}(F_i)=(F_1|\cdots|F_{i-1}|F_{i+1}|\cdots|F_s)$, $\mathrm{complem}(G_i)=(G_1|\cdots|G_{i-1}|G_{i+1}|\cdots|G_s)$ *and* $\mathrm{sep}_F(B,\ C)=\min_{\|Q\|_F=1}\|QB-CQ\|_F$ *denotes the separation of two square matrices B and*

C. Let E be a perturbation matrix. Then there is an $\varepsilon>0$ such that if $\|E\|_F<\varepsilon$, there exist matrices $A'_{\lambda_i}\in\mathbb{C}^{r_i\times r_i}$, $i=1, 2, \cdots, s$, with column full rank matrices $G'_i\in\mathbb{C}^{n\times r_i}$ such that

$$(A+E)(G'_1|G'_2|\cdots|G'_s)=(G'_1|G'_2|\cdots|G'_s)\operatorname{diag}[A'_1, A'_2, \cdots, A'_s], \tag{1}$$

where the matrix $(G'_1|G'_2|\cdots|G'_s)$ is nonsingular, and the inequalities

$$\|A_{\lambda_i}-A'_i\|_F\leqslant k_i\|E\|_F+\frac{2k_i\cdot k'_i}{s_i-(k_i+k'_i)\|E\|_F}\|E\|_F^2, \tag{2}$$

$$\|G_i-G'_i\|_F\leqslant\frac{2k_i\cdot k'_i}{\|F_i\|_F\cdot[s_i-(k_i+k'_i)\|E\|_F]}\|E\|_F \tag{3}$$

hold for $i=1, 2, \cdots, s$.

Proof Write $F_{ij}=F_i^HEG_j$, $F=(F_{ij})_{s\times s}$. The complementary submatrix of $A_{\lambda_i}+F_{ii}$ in $\operatorname{diag}[A_{\lambda_1}, A_{\lambda_2}, \cdots, A_{\lambda_s}]+F$ is denoted by $\operatorname{complem}(A_{\lambda_i}+F_{ii})$.

When

$$\frac{\|(F_{i1}|\cdots|F_{ii-1}|F_{ii+1}|\cdots|F_{is})\|_F\cdot\|(F_{1i}^H|\cdots|F_{i-1i}^H|F_{i+1i}^H|\cdots|F_{si}^H)\|_F}{\operatorname{sep}_F^2(A_{\lambda_i}+F_{ii}, \operatorname{complem}(A_{\lambda_i}+F_{ii}))}<\frac{1}{4}, \tag{4}$$

the matrix equation

$$\begin{aligned}&P(A_{\lambda_i}+F_{ii})-(\operatorname{complem}(A_{\lambda_i}+F_{ii}))P\\&\quad=(F_{1i}^H|\cdots|F_{i-1i}^H|F_{i+1i}^H|\cdots|F_{si}^H)^H-P(F_{i1}|\cdots|F_{ii-1}|F_{ii+1}|\cdots|F_{is})P\end{aligned}$$

has a solution P_i satisfying

$$\|P_i\|_F\leqslant\frac{2\|(F_{1i}^H|\cdots|F_{i-1i}^H|F_{i+1i}^H|\cdots|F_{si}^H)\|_F}{\operatorname{sep}_F(A_{\lambda_i}+F_{ii}, \operatorname{complem}(A_{\lambda_i}+F_{ii}))}. \tag{5}$$

Write $A'_i=A_{\lambda_i}+F_{ii}+(F_{i1}|\cdots|F_{ii-1}|F_{ii+1}|\cdots|F_{is})P_i$. Then

$$P_{\sigma_i}^T(\operatorname{diag}[A_{\lambda_1}, A_{\lambda_2}, \cdots, A_{\lambda_s}]+F)P_{\sigma_i}\begin{pmatrix}I\\P_i\end{pmatrix}=\begin{pmatrix}I\\P_i\end{pmatrix}A'_i, \tag{6}$$

where $P_{\sigma_i}=(e_{\sigma_i(1)}|e_{\sigma_i(2)}|\cdots|e_{\sigma_i(n)})$ is a permutation matrix and

$$\sigma_i=\begin{pmatrix}1, & 2, \cdots, u_1, u_1+1, \cdots, u_2, \cdots\cdots & \\ u_{i-1}+1, \cdots, u_i, & 1, \cdots, u_1, \cdots, u_{i-2}+1, \cdots, u_{i-1}, u_{i+1}, \cdots, u_{i+1}, & \cdots\end{pmatrix},$$

$$u_i=\sum_{t=1}^{i}r_t$$

is a permutation. Let

$$G'_i=(G_1|G_2|\cdots|G_s)P_{\sigma_i}\begin{pmatrix}I\\P_i\end{pmatrix}. \tag{7}$$

Then by (6) we have

$$(A+E)(G'_1|G'_2|\cdots|G'_s)=(G'_1|G'_2|\cdots|G'_s)\operatorname{diag}[A'_1, A'_2, \cdots, A'_s],$$

and the inequalities

$$\|A_{\lambda_i}-A'_i\|_F\leqslant\|F_{ii}\|_F+\|(F_{i1}|\cdots|F_{ii-1}|F_{ii+1}|\cdots|F_{is})P_i\|_F, \tag{8}$$

$$\|G_i-G'_i\|_F\leqslant\|P_i\|_F\cdot\|(G_1|\cdots|G_{i-1}|G_{i+1}|\cdots|G_s)\|_F \tag{9}$$

hold.

The matrix $\left(P_{\sigma_1}\begin{pmatrix}I\\P_1\end{pmatrix}\Big|P_{\sigma_2}\begin{pmatrix}I\\P_2\end{pmatrix}\Big|\cdots\Big|P_{\sigma_s}\begin{pmatrix}I\\P_s\end{pmatrix}\right)$ is nonsingular if

$$\sum_{i=1}^{s}\|P_i\|_F^2<1. \tag{10}$$

By 2° $(G_1|G_2|\cdots|G_s)$ is nonsingular. Hence $(G'_1|G'_2|\cdots|G'_s)$ is nonsingular.

By the properties of separation, inequalities

$$\begin{aligned}&\operatorname{sep}_F(A_{\lambda_i}+F_{ii},\ \operatorname{complem}(A_{\lambda_i}+F_{ii}))\\ &\geqslant \operatorname{sep}_F(A_{\lambda_i},\ \operatorname{diag}[A_{\lambda_1},\ \cdots,\ A_{\lambda_{i-1}},\ A_{\lambda_{i+1}},\ \cdots,\ A_{\lambda_s}])-\|F_{ii}\|_F-\|\operatorname{complem}(F_{ii})\|_F\\ &\geqslant \min_{j\neq i}\operatorname{sep}_F(A_{\lambda_i},\ A_{\lambda_j})-(\|F_i\|_F\cdot\|G_i\|_F+\|\operatorname{complem}(F_i)\|_F\cdot\|\operatorname{complem}(G_i)\|_F)\|E\|_F\end{aligned} \tag{11}$$

hold.

The inequalities (4) and (10) are valid for sufficiently small $\|E\|_F$. From (5), (11), (8) and (9), the result of Theorem 1 is obtained.

The result of Theorem 1 must be understood as follows.

Under a small perturbation, a multiple eigenvalue of a matrix is generally split into a group of the eigenvalues near it. Therefore, the eigenvalue subspace will be completely changed, and it possesses continuity only in the sense of the invariant subspace.

Suppose that the multiple eigenvalue λ_i of A of multiplicity r_i continuously changed to the eigenvalues of $A+E$. Then for sufficiently small $\|E\|_F$, they form a group which is separated from other groups of the eigenvalues of $A+E$. Write this group by Λ_i. Then the Λ_i defines a subspace V_i. It is direct sum of all right eigenvalue subspaces of the eigenvalues of Λ_i.

Let there be $A_i\in\mathbf{C}^{r_i\times r_i}$ and column full rank matrix H_i such that $(A+E)H_i=H_iA_i$, and all of the eigenvalues of A_i form Λ_i. Then by 2° of Section 1, there exists nonsingular matrix T_i such that $T_i^{-1}A_iT_i$ is a block diagonal matrix whose diagonal blocks are all the eigenvalue matrices of A_i. Since other $n-r_i$ eigenvalues of $A+E$ differ from those of A_i, so that those diagonal blocks are also eigenvalue matrices of $A+E$ and the columns of H_iT_i also form a basis of V_i. Hence the columns of H_i form a basis of V_i. This fact implies that the V_i is a unique invariant subspace defined by Λ_i.

By general definition in [8], the distance between subspace $N((A-\lambda_iI)^{r_i})$ and subspace V_i is $d_F(G_i,\ G'_i)=\{r_i-\operatorname{tr}[(G_i^HG_i)^{-1}G_i^HG'_i(G'^H_iG'_i)^{-1}G'^H_iG_i]\}^{\frac{1}{2}}$. Using(7), We have $d_F(G_i,\ G'_i)=O(\|P_i\|_F)$. Hence by (5), $d_F(G_i,\ G'_i)=O(\|E\|_F)$. This shows how the eigenvalue subspace $N((A-\lambda_iI)^{\lambda_i})$ is continually changed to V_i as invariant subspace.

Theorem 1 denotes that there is a column full rank matrix G'_i near G_i, it is formed by the vectors of a basis of V_i and determines a matrix A'_i near A_i. Moreover, the set of the eigenvalues of $A+E$ is union of the sets of the eigenvalues of A'_i $(i=1,\ 2,\ \cdots,\ s)$. In such sense, (2) and (3) can be regaded as a generalization of the sensitivity of the simple eigenvalue and the corresponding eigenvector. The scalar k_i determined by a dual basis pair of the eigenvalue subspaces is a generalization of Wilkinson's condition number.

3. Sensitivity of the multiple eigenvalue

In [1], Golub and Wilkinson indicated the following fact. For the quasi-Jordan block

$$\tilde{J}_1=\begin{pmatrix} 1 & \theta_1 & & & \\ & 1 & \theta_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \theta_{r-1} \\ & & & & 1 \end{pmatrix}_{r\times r},$$

the eigenvalue 1 is highly sensitive when $\theta_1=\theta_2=\cdots=\theta_{r-1}=1$, and its perturbation can be of $\|E\|_2^{\frac{1}{r}}$. But if $\theta_1=\theta_2=\cdots=\theta_{r-1}=10^{-10}$, then, since the perturbation is of order 10^{-10} when $\|E\|_2=10^{-10}$, so the eigenvalue 1 is not at all sensitive.

For general matrix, the recent estimate which is obtained by Kahan, Parlett and Jiang by means of the classical Jordan canonical form is as follows.

(*) Let $J=T^{-1}AT$ be A's Jordan form and let E be a perturbation matrix. To any eigenvalue λ of $A+E$ there corresponds an eigenvalue λ_i of A such that

$$\frac{|\lambda-\lambda_i|^{l_i}}{(1+|\lambda-\lambda_i|)^{l_i-1}}\leqslant\|T^{-1}\|\cdot\|T\|\cdot\|E\|,$$

where l_i is the order of the largest Jordan block to which λ_i belongs. The specrtal (or Frobenius) norm must be used here.

Use (*) on $\tilde{J}_1$. In the first case, we have estimate $\frac{|\lambda-1|^r}{(1+|\lambda-1|)^{r-1}}\leqslant\|E\|_2$. But in the second case, this estimate is $\frac{|\lambda-1|^r}{(1+|\lambda-1|)^{r-1}}\leqslant\|D^{-1}\|_2\cdot\|D\|_2\cdot\|E\|_2$, where D is a diagonal matrix. The perturbation bound of the latter is increased contrarily, so the behavior of $\tilde{J}_1$ under perturbations can not be interpreted by (*).

Theorem 2. *Let $A_\lambda\in\mathbf{C}^{r\times r}$ be an eigenvalue matrix and let*

$$\tilde{J}_\lambda=\begin{pmatrix} \lambda & \theta_1 & & & \\ & \lambda & \theta_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \theta_{r-1} \\ & & & & \lambda \end{pmatrix}_{r\times r}=T^{-1}A_\lambda T$$

be a quasi-Jordan matrix of A_λ, where T is a nonsingular matrix. Then, for a perturbation matrix E, any eigenvalue λ' of $A_\lambda+E$ satisfies inequality

$$\frac{|\lambda'-\lambda|^r}{[|\lambda'-\lambda|^{2(r-1)}+\delta_{r-2}|\lambda'-\lambda|^{2(r-2)}+\cdots+\delta_1|\lambda'-\lambda|^2+\delta_0]^{\frac{1}{2}}}\leqslant\alpha\cdot\|T^{-1}\|_2\cdot\|T\|_2\cdot\|E\|_2,$$

where $\delta_t=\frac{1}{r}\sum_{h=1}^{t+1}|\theta_h\theta_{h+1}\cdots\theta_{h+r-t-2}|^2,\ t=0,\ 1,\ \cdots,\ r-2,\ 1\leqslant\alpha\leqslant\sqrt{r}$.

Corollary. *Under the hypotheses of Theorem 2, for any fixed* Δ *and* δ *satisfying* $0<\delta<\Delta$, *there is an* $\varepsilon>0$ *such that if* $\|E\|_2<\varepsilon$, *any eigenvalue* λ' *of* $A+E$ *which does not lie in the disk* $|z-\lambda|<\delta$ *lies in the annulus*

$$\delta\leqslant|z-\lambda|\leqslant[\alpha\cdot C\cdot\|T^{-1}\|_2\cdot\|T\|_2\cdot\|E\|_2]^{\frac{1}{m}},$$

where m *is the minimum of natural number* e *such that for* $t\geqslant e$, *the inequality* $r\delta_{r-t-1}<1$ *is valed,* $C=\left[\Delta^{2(m-1)}+\delta_{r-2}\Delta^{2(m-2)}+\cdots+\delta_{r-m}+\frac{\delta_{r-m-1}}{\delta^2}+\cdots+\frac{\delta_0}{\delta^{2(r-m)}}\right]^{\frac{1}{2}}$, $\delta_{r-1}=1,\ \delta_{-1}=0$.

Moreover, $m\leqslant l$, *where* l *is the order of the largest Jordan block of* A_λ.

Proof Let λ' be an eigenvalue of $A_\lambda+E$ and let $\lambda'-\lambda=a\neq0$. Since the matrix

$$I-(\lambda'I-\tilde{J}_\lambda)^{-1}T^{-1}ET=(\lambda'I-\tilde{J}_\lambda)^{-1}[\lambda'I-\tilde{J}_\lambda-T^{-1}ET]$$

is singular, so that

$$1\leqslant\|(\lambda'I-\tilde{J}_\lambda)^{-1}T^{-1}ET\|_2\leqslant\|(\lambda'I-\tilde{J}_\lambda)^{-1}\|_2\cdot\|T^{-1}\|_2\cdot\|T\|_2\cdot\|E\|_2.$$

However, $\|(\lambda'I-\tilde{J}_\lambda)^{-1}\|_F=\frac{\sqrt{r}}{|a|}\left[\sum_{t=0}^{r-1}\delta_t|a|^{2t}\right]^{\frac{1}{2}}$ and $\frac{1}{\sqrt{r}}\|\cdot\|_F\leqslant\|\cdot\|_2\leqslant\|\cdot\|_F$. Hence the result of this theorem is obtained.

For the sufficiently small $\|E\|_2$, $|a|\leqslant\Delta$. Take λ' such that $|a|\geqslant\delta$, then

$$|a|^m\leqslant\alpha\cdot\|T^{-1}\|_2\cdot\|T\|_2\cdot\|E\|_2\cdot[|a|^{2(m-1)}+\delta_{r-2}|a|^{2(m-2)}+\cdots$$
$$+\delta_{r-m}+\frac{\delta_{r-m-1}}{|a|^2}+\cdots+\frac{\delta_0}{|a|^{2(r-m)}}\Bigg]^{\frac{1}{2}}.$$

Taking m as the natural number given in Corollray, the Corollary is also obtained.

The number $r\delta_{r-t-1}$ is the sum of the square of the absolute value of the tth superdiagonal elements of the matrix

$$A_{(\lambda)}=\begin{pmatrix}1 & \theta_1 & \theta_1\theta_2\cdots\cdots & \theta_1\theta_2\cdots\theta_{r-1}\\ & 1 & \theta_2\cdots\cdots & \theta_2\cdots\theta_{r-1}\\ & & 1\ \ddots & \vdots\\ & & \ddots\ \ddots & \vdots\\ & & \ddots & \theta_{r-1}\\ & & & 1\end{pmatrix}.$$

For some $\theta_t=0$, the $A_{(\lambda)}$ is split into some blocks in conformity to the Jordan form of A_λ, and the lth, $l+1$th, $\cdots$, $r-1$th superdiagonal elements are all zero. Hence $m<l$.

In disk $|z-\lambda|<\delta$, Corollary gives no description on the sensitivity. However, δ and Δ are quite arbitrary, so that Corollary has reasonably described the perturbation of λ yet. In fact, there is no need to study very small change of λ.

Consider $\tilde{J}_1$ again. Let $r=10$ and $\|E\|_F=10^{-10}$. Since $|e_{ij}|\leqslant\|E\|_2$ ($E=(e_{ij})$), the inequality $|\alpha|\leqslant 1+10\|E\|_2=1+10^{-9}$ can be deduced by Gerschgorin theorem. Take $\Delta=1+10^{-9}$. Then for any δ satisfying $0<\delta<\Delta$, $C<\sqrt{5.6}$. Hence the annulus has turned into a disk $|z-1|\leqslant(\alpha\cdot 5.6^{\frac{1}{2}}\cdot 10^{-10})^{\frac{1}{10}}$. So there is an estimate $|\lambda-1|<0.1223$. It should be observed that this is a good estimate. If $\theta_1=\theta_2=\cdots=\theta_{r-1}=10^{-10}$, then $\delta_t=\frac{t+1}{10}10^{-20(9-t)}$. So $m=1$. Take $\delta=10^{-10}$. Then for any Δ satisfying $10^{-10}<\Delta$, $C\leqslant\sqrt{5.5}$. Hence the annulus is $10^{-10}\leqslant|z-1|\leqslant\alpha\cdot 5.5^{\frac{1}{2}}\cdot 10^{-10}$. Since $\sqrt{5.5}\doteq 2.4$, and in this case, $\alpha\doteq 1$, So we know that the bound of perturbation is of order 10^{-10} really.

Thus, the behavior of the eigenvalue of $\tilde{J}_1$ is interpreted quantitatively.

4. The Perturbation theorem

Note that the separation $\mathrm{sep}(B, C)$ of two square matrices B and C is the smallest singular value of $I\otimes B-C^T\otimes I$, where the Kronecker product of X and Y is denoted by $X\otimes Y$, and note that there is some natural number i_0 such that $\|\mathrm{diag}[(\lambda I-\tilde{J}_{\lambda_1})^{-1}, \cdots, (\lambda I-\tilde{J}_{\lambda_s})^{-1}\|_2=\|(\lambda I-\tilde{J}_{\lambda_{i_0}})^{-1}\|_2$. Then by Theorems 1 and 2, we can obtain the main theorem as follows.

Theorem 3. *Let $A\in\mathbf{C}^{n\times n}$. Let $\tilde{J}=\mathrm{diag}[\tilde{J}_{\lambda_1}, \tilde{J}_{\lambda_2}, \cdots, \tilde{J}_{\lambda_s}]=T^{-1}AT$ be a quasi-Jordan matrix of A, where $\tilde{J}_{\lambda_i}=\lambda_i I+\tilde{J}_i$, all λ_i are different, and*

$$\tilde{J}_i=\begin{pmatrix} 0 & \theta_1^{(i)} & & & \\ & 0 & \theta_2^{(i)} & & \\ & & \ddots & \ddots & \\ & & & \ddots & \theta_{r_i-1}^{(i)} \\ & & & & 0 \end{pmatrix}_{r_i\times r_i}.$$

Let $T=(G_1|G_2|\cdots|G_s)$, $T^{-1}=(F_1|F_2|\cdots|F_s)^H$, where G_i, $F_i\in\mathbf{C}^{n\times r_i}$. For $i=1, 2, \cdots, s$, set $k_i=\|G_i\|_F\cdot\|F_i\|_F$, $k_i'=\|\mathrm{complem}(G_i)\|_F\cdot\|\mathrm{complem}(F_i)\|_F$, $\Sigma_i=\min_{j\neq i}\sigma_{\min}(T_{ij})$, where complem $(G_i)=(G_1|\cdots|G_{i-1}|G_{i+1}|\cdots|G_s)$; complem$(F_i)=(F_1|\cdots|F_{i-1}|F_{i+1}|\cdots|F_s)$, *$\sigma_{\min}(T_{ij})$ is the smallest singular value of T_{ij}, and $T_{ij}=(\lambda_i-\lambda_j)I+I\otimes\tilde{J}_i-\tilde{J}_j\otimes I$.*

Let E be a perturbation matrix. Then there is an $\varepsilon>0$ such that if $\|E\|_F<\varepsilon$, any eigenvalue λ of $A+E$ satisfies one of the inequalities

$$\frac{|\lambda-\lambda_i\|^{r_i}}{\alpha_i[|\lambda-\lambda_i|^{2(r_i-1)}+\delta_{r_i-2}^{(i)}|\lambda-\lambda_i|^{2(r_i-2)}+\cdots+\delta_1^{(i)}|\lambda-\lambda_i|^2+\delta_0^{(i)}]^{\frac{1}{2}}}\leqslant R_i^{(k)}(E),\ i=1, 2, \cdots, s,$$

where

$$R_i^{(1)}(E)=\|T^{-1}\|_2\cdot\|T\|_2\cdot\|E\|_2,$$

$$R_i^{(2)}(E)=\beta_i\left(k_i\|E\|_F+\frac{2k_i\cdot k_i'}{\Sigma_i-(k_i+k_i')\|E\|_F}\|E\|_F^2\right),\quad 1\leqslant\alpha_i\leqslant\sqrt{r_i},$$

$$\frac{1}{\sqrt{r_i}}\leqslant\beta_i\leqslant 1,\ \delta_j^{(i)}=\sum_{h=1}^{j+1}|\theta_h^{(i)}\theta_{h+1}^{(i)}\cdots\theta_{h+r_i-j-2}^{(i)}|^2,\ \delta_{r_i-1}^{(i)}=1,\ j=0, 1, \cdots, r_i-2, \textit{and } k=1 \textit{ or } 2.$$

Corollary. *Under the hypotheses of Theorem 3, for any fixed Δ and δ satisfying $0<\delta<\Delta$, there is an $\varepsilon>0$ such that if $\|E\|_F<\varepsilon$, any eigenvalue λ of $A+E$, which does not lie in any disk $|z-\lambda_i|<\delta$, lies in the union of the annuluses*

$$\delta\leqslant|z-\lambda_i|\leqslant[\alpha_i\cdot C_i\cdot R_i^{(k)}(E)]^{\frac{1}{m_i}},\ i=1,\ 2,\ \cdots,\ s,$$

where m_i is the minimum of the natural number e such that for $t\geqslant e$, $r_i\delta^{(i)}_{r_i-t-1}<1$ is valid

$$c_i=\left[\Delta^{2(m_i-1)}+\delta^{(i)}_{r_i-2}\Delta^{2(m_i-2)}+\cdots+\delta^{(i)}_{r_i-m_i}+\frac{\delta^{(i)}_{r_i-m_i-1}}{\delta^2}+\cdots+\frac{\delta^{(i)}_0}{\delta^{2(r_i-m_i)}}\right]^{\frac{1}{2}},\ \delta^{(i)}_{-1}=0.$$

Moreover, $m_i\leqslant l_i$, $i=1,\ 2,\ \cdots,\ s$, where l_i is the order of the largest Jordan block to which λ_i belongs.

For a sufficint small $\|E\|_F$, the annuluses in this Corollary do not overlap each othre. So, for the study of perturbation of λ_i, we consider the i-th annulus only. But for the relatively large $\|E\|_F$, those annuluses are overlaping each other.

When λ_i is simple for some i, we have $r_i=1$, $\alpha_i=1$ and $\beta_i=1$. Taking $k=2$, Theorem 3 gives estimate $|\lambda-\lambda_i|\leqslant k_i\|E\|_F+\dfrac{2k_i\cdot k_i'}{\Sigma_i-(k_i+k_i')\|E\|_F}\|E\|_F^2$. It is essentially consistent with the resuls of Wilkinson and Stewart, where k_i is Wilkinson's condition number. In the case that A is a nondefective matrix, $\Sigma_i=\min\limits_{j\neq i}|\lambda_i-\lambda|$.

When A is a nondefective matrix, $\alpha_i=1$, $m_i=1$ and $c_i=1$, for $i=1,\ 2,\ \cdots,\ s$. Taking $k=1$, Theorem 3 and its Corollary give the result of Bauer and Fike, i. e.

$$|\lambda-\lambda_i|\leqslant\|T^{-1}\|_2\cdot\|T\|_2\cdot\|E\|_2.$$

Since the dual basis pair producing $\tilde{J}_{\lambda_i}$ is not unique, so that k_i is not unique for $\tilde{J}_{\lambda_i}$. On the other hand, when each k_i is a minimum for $\tilde{J}_{\lambda_i}$, the J is a "natural" quasi-Jordan matrix of A for computational purposes. Therefore, as an example, for $\tilde{J}_1$ in Section 3, we must take it as $\tilde{J}$ of Theorem 3 but not as $D^{-1}\tilde{J}_1D$, where D is a diagonal matrix $(D\neq I)$.

In brief, the sensitivity of the eigenvalues of a defective matrix (or a general matrix) depends on distribution of its eigenvalues, on a dual basis pair of each eigenvalue subspace pair and on superdiagonal elements of the quasi-Jordan form corresponding to this dual basis pair such that k_i is a minimum. This is just a point of view Golub and Wilkinson attempted to show in [1].

References

[1] Golub, G. H., Wilkinson, J. H., *SIAM Rev.*, **18** (1976).
[2] Stewart, G. W., *SIAM Rev.*, **15** (1973).
[3] Varah, J. M., On separation of two metrices, *SIAM J. Numer. Anal.*, **16** (1979).
[4] Wilkinson, J. H., London and New York. (1965).
[5] Stewart, G. W., Introduction to Matrix Computation, Academic Press, (1973).
[6] Bauer, F. L., Fike, C. T., Norms and exclusion theorem, *Numer. Math.*, **2** (1960).
[7] Kahan, W., Parlett, B. N., Jiang, E. X., *SIAM J. Numer. Anal.*, **19** (1982).
[8] Sun Ji-guang, *Mathematica Numerica Sinica*, **2** (1980).