Improving the condition number of a simple eigenvalue by a rank one matrix ☆

Rafael Bru[a], Rafael Cantó[a,*], Ana M. Urbano[a]

[a]*Institut de Matemàtica Multidisciplinar, Universitat Politècnica de València, 46071 València, Spain.*

Abstract

In this work a technique to improve the condition number s_i of a simple eigenvalue λ_i of a matrix $A \in \mathbb{C}^{n\times n}$ is given. This technique obtains a rank one updated matrix that is similar to A with the eigenvalue condition number of λ_i equal to one. More precisely, the similar updated matrix $A + v_i q^*$, where $Av_i = \lambda_i v_i$ and q is a fixed vector, has $s_i = 1$ and the remaining condition numbers are at most equal to the corresponding initial condition numbers. Moreover an expression to compute the vector q, using only the eigenvalue λ_i and its eigenvector v_i, is given.

Keywords: Brauer's theorem, diagonalization matrix, eigenvalue condition number, rank-one update.
AMS classification: 15A18, 15A23

1. Introduction

Let $A \in \mathbb{C}^{n\times n}$ and let λ_i be a simple eigenvalue of A with associated right and left eigenvectors v_i and l_i, respectively. The *condition number* of λ_i is given by

$$s_i = \frac{\|v_i\|\,\|l_i\|}{|l_i^* v_i|} \geq 1,$$

that is, s_i is the inverse of the cosine of the angle between the right and left eigenvectors of A associated with λ_i (see [6, 8, 9]). To compute s_i some authors assume that the right and left eigenvectors are normalized. However, we assume that the right eigenvectors are normalized and the left eigenvectors are chosen in such away that $l_i^* v_i = 1$.

The interpretation of the condition number of an eigenvalue λ_i is that an $\mathcal{O}(\epsilon)$ perturbation in A can cause an $\mathcal{O}(\epsilon s_i)$ perturbation in the eigenvalue λ_i. So, if s_i is near to 1 a perturbation in A will have less effect. Byers and Kressner

☆Supported by the Spanish DGI grant MTM2013-43678-P
e-mail: rbru@mat.upv.es, rcanto@mat.upv.es, amurbano@mat.upv.es
*Corresponding author

[5] study the variation of the condition number of a complex eigenvalue under a real perturbation and they show that restricting the backward error to be real the condition number decreases at most by a factor of $1/\sqrt{2}$. Therefore, an interesting and more general problem is the following: Can we update the matrix A maintaining the same spectrum and improving the corresponding eigenvalue condition numbers?

In this work, we show that an $n \times n$ complex matrix with n distinct and ill conditioned eigenvalues can be updated, with a rank one perturbation, to a similar matrix such that one of its eigenvalue condition number is one and the remaining eigenvalue condition numbers are less or equal than the corresponding of those of the matrix A. In addition, the sensitivity of eigenvectors are given. Finally, Theorem 2 gives a method to obtain this rank one perturbation where it is only necessary to know one eigenvalue and its corresponding right eigenvector.

It is worth to note that the rank one modification has also been used to update the singular value decomposition [3] and the symmetric eigenproblem [4].

2. Improving eigenvalue condition numbers

In this section we apply the Brauer's Theorem and the results given in [2, 7] to improve the eigenvalue condition number of a matrix with pairwise distinct eigenvalues.

Theorem 1. *Let $A \in \mathbb{C}^{n\times n}$ be a matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, pairwise distinct, and $v_1, v_2, \ldots, v_n$, their associated unit right eigenvectors. Let $s_1, s_2, \ldots, s_n$, be the corresponding eigenvalue condition numbers. Then, there exists an n-dimensional vector $q_{(1)}$, with $q_{(1)}^* v_1 = 0$, such that the matrix $A^{(1)} = A + v_1 q_{(1)}^*$ is similar to A and the corresponding condition numbers of its eigenvalues satisfy that $s_1^{(1)} = 1$ and $s_i^{(1)} \le s_i$, for $i = 2, 3, \ldots, n$. Moreover, if $v_1^{(1)}, v_2^{(1)}, \ldots, v_n^{(1)}$ are the associated eigenvectors of $A^{(1)}$, then*

$$\left\| v_i^{(1)} - v_i \right\| = | < v_i, v_1 > | = |v_1^* v_i| \quad i = 2, 3, \ldots, n.$$

Proof. Let q be an arbitrary solution of the equation $q^* v_1 = 0$. By the Brauer's Theorem (see [1, 2]) A and $A + v_1 q^*$ are similar matrices. Let $l_1, l_2, \ldots, l_n$, be the left eigenvectors of A associated with $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively, and such that $l_j^* v_i = \delta_{ij}$, $i, j = 1, 2, \ldots, n$. Then, the eigenvalue condition numbers are

$$s_i = \frac{\|v_i\| \, \|l_i\|}{|l_i^* v_i|} = \|l_i\|, \quad i = 1, 2, \ldots, n.$$

By [2, Propositions 1.1. and 1.2.] and [7], the right $\{w_1, w_2, \ldots, w_n\}$ and left $\{r_1, r_2, \ldots, r_n\}$ eigenvectors of $A + v_1 q^*$ associated with $\lambda_1, \lambda_2, \ldots, \lambda_n$, are respectively

$$\left.\begin{array}{lll} w_1 = v_1, & w_i = v_i - \dfrac{q^* v_i}{\lambda_1 - \lambda_i}\, v_1, & i = 2, 3, \ldots, n, \\ r_1^* = l_1^* + \sum_{i=2}^{n} \dfrac{q^* v_i}{\lambda_1 - \lambda_i}\, l_i^*, & r_i^* = l_i^*, & i = 2, 3, \ldots, n. \end{array}\right\} \quad (1)$$

Since

$$\begin{aligned} r_1^* w_1 &= r_1^* v_1 = \left(l_1^* + \sum_{i=2}^{n} \frac{q^* v_i}{\lambda_1 - \lambda_i}\, l_i^* \right) v_1 = l_1^* v_1 + \sum_{i=2}^{n} \frac{q^* v_i}{\lambda_1 - \lambda_i}\, l_i^* v_1 = 1, \qquad (2) \\ r_i^* w_i &= l_i^* w_i = l_i^* \left(v_i - \frac{q^* v_i}{\lambda_1 - \lambda_i}\, v_1 \right) = l_i^* v_i - \frac{q^* v_i}{\lambda_1 - \lambda_i}\, l_i^* v_1 = 1, \quad i = 2, 3, \ldots, n, \end{aligned}$$

the condition numbers $\tilde{s}_i$ of the eigenvalues λ_i, $i = 1, 2, \ldots, n$, of the updated matrix $A + v_1 q^*$ are

$$\left.\begin{aligned} \tilde{s}_1 &= \frac{\|w_1\|\, \|r_1\|}{|r_1^* w_1|} = \|v_1\|\, \|r_1\| = \|r_1\|, \\ \tilde{s}_i &= \frac{\|w_i\|\, \|r_i\|}{|r_i^* w_i|} = \|w_i\|\, \|l_i\| = \|w_i\|\, s_i, \quad i = 2, 3, \ldots, n. \end{aligned}\right\} \quad (3)$$

Therefore, $\tilde{s}_i \leq s_i$, whenever $\|w_i\| \leq 1$, for $i = 2, 3, \ldots, n$.
Since $w_i = v_i - \dfrac{q^* v_i}{\lambda_1 - \lambda_i}\, v_1$, by the approximation theory the vector w_i has minimal norm when $\dfrac{q^* v_i}{\lambda_1 - \lambda_i}\, v_1$ is the orthogonal projection of v_i on span$\{v_1\}$, that is, when

$$\frac{q^* v_i}{\lambda_1 - \lambda_i}\, v_1 = \mathrm{Proj}_{v_1}(v_i) = \frac{< v_i,\, v_1 >}{\|v_1\|^2}\, v_1 = (v_1^* v_i)\, v_1.$$

Then, we need that the vector q satisfies the following system

$$\left.\begin{aligned} q^* v_1 &= 0, \\ q^* v_i &= (\lambda_1 - \lambda_i)\, (v_1^* v_i)\,, \ i = 2, 3, \ldots, n. \end{aligned}\right\} \quad (4)$$

Let $q_{(1)}$ be the unique solution of this consistent system. Consider now the updated matrix with this unique solution $A^{(1)} = A + v_1 q_{(1)}^*$ and let us denote the eigenvectors of this matrix with the superscript (1). By (1) the right and left eigenvectors of $A^{(1)}$, $\{v_1^{(1)}, v_2^{(1)}, \ldots, v_n^{(1)}\}$ and $\{l_1^{(1)}, l_2^{(1)}, \ldots, l_n^{(1)}\}$ respectively, associated with $\lambda_1, \lambda_2, \ldots, \lambda_n$, are given by

$$\left.\begin{array}{lll} v_1^{(1)} = v_1, & v_i^{(1)} = v_i - (v_1^* v_i)\, v_1, & i = 2, 3, \ldots, n, \\ \left(l_1^{(1)}\right)^* = l_1^* + \sum_{i=2}^{n} (v_1^* v_i)\; l_i^*, & \left(l_i^{(1)}\right)^* = l_i^*, & i = 2, 3, \ldots, n. \end{array}\right\} \quad (5)$$

Since v_i and v_1 are unit vectors, note that $\left\| v_i^{(1)} \right\| \leq 1$, for $i = 2, 3, \ldots, n$. Then, by equation (3) applied to the right eigenvector $v_i^{(1)}$, the corresponging eigenvalue condition numbers of $A^{(1)}$ satisfy

$$s_i^{(1)} = \left\| v_i^{(1)} \right\| \, s_i \leq s_i, \quad i = 2, 3, \ldots, n.$$

It remains to prove that $s_1^{(1)} = \left\| l_1^{(1)} \right\| = 1$. The right and left eigenvectors of $A^{(1)}$ satisfy

$$\begin{aligned} \langle v_i^{(1)}, v_1^{(1)} \rangle &= \left(v_1^{(1)} \right)^* v_i^{(1)} = 0, \text{ and} \\ < v_i^{(1)}, l_1^{(1)} >= \left(l_1^{(1)} \right)^* v_i^{(1)} &= \left(l_1^* + \sum_{j=2}^{n} (v_1^* v_j) \, l_j^* \right) (v_i - (v_1^* v_i) \, v_1) = 0, \end{aligned}$$

for $i = 2, 3, \ldots n$. Then

$$l_1^{(1)} \in \text{ span} \left\{ v_2^{(1)}, \, v_3^{(1)}, \, \ldots, \, v_{n-1}^{(1)}, \, v_n^{(1)} \right\}^{\perp} = \text{span} \left\{ v_1^{(1)} \right\},$$

and therefore

$$l_1^{(1)} = \alpha \, v_1^{(1)}.$$

Applying equation (2) to the new eigenvectors we have $\left(l_1^{(1)} \right)^* v_1^{(1)} = 1$. On the other hand,

$$\left(l_1^{(1)} \right)^* v_1^{(1)} = \left(\overline{\alpha} \left(v_1^{(1)} \right)^* \right) v_1^{(1)} = \overline{\alpha} \left\| v_1^{(1)} \right\|^2 = \overline{\alpha}.$$

Then, $\alpha = 1$ and

$$l_1^{(1)} = v_1^{(1)}. \tag{6}$$

Using equation (3) $s_1^{(1)} = 1$, since $\left\| v_1^{(1)} \right\| = 1$.
Finally, by (5) we obtain

$$\left\| v_i^{(1)} - v_i \right\| = | < v_i, \, v_1 > | = |v_1^* v_i| \quad i = 2, 3, \ldots, n.$$

□

We illustrate the results of Theorems 1 with the following example, where we have used MatLab.

Example 1. Consider the matrix

$$A = \begin{bmatrix} -149 & -50 & -154 & -1 \\ 537 & 180 & 546 & 2 \\ -27 & -9 & -25 & 1 \\ 0 & 0 & 0 & 2.9999 \end{bmatrix}$$

with eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $\lambda_4 = 2.9999$, and the corresponding eigenvalue condition numbers

$$\begin{aligned} s_1 &= 619.826169515, & s_2 &= 437.718033299, \\ s_3 &= 1006180.948310136, & s_4 &= 1006143.406357263. \end{aligned}$$

Applying Theorem 1 with the right eigenvector v_1 associated with $\lambda_1 = 1$ we obtain the matrix $A^{(1)} = A + v_1 q_{(1)}^*$, similar to A, such that its eigenvalue condition numbers are

$$\begin{array}{rclrcl} s_1^{(1)} & = & 1, & s_2^{(1)} & = & 60.9478235921, \\ s_3^{(1)} & = & 252507.4326370870, & s_4^{(1)} & = & 252533.6146298055. \end{array}$$

Remark 1. Note that if we apply Theorem 1 to the matrix of the Example 1 using the eigenvalue $\lambda_3 = 3$ we obtain the updated matrix $A^{(1)} = A + v_3 q_{(3)}^*$ with the eigenvalue condition numbers

$$\begin{array}{rclrcl} s_1^{(1)} & = & 155.5492761525672, & s_2^{(1)} & = & 167.7693394271733, \\ s_3^{(1)} & = & 1, & s_4^{(1)} & = & 36.7972549222763. \end{array}$$

This fact shows that the improvement of the eigenvalue condition numbers depends on the eigenvector with we are working on. Then, to choose the eigenvector to use is a natural question. The following theorem gives some insight on this question.

Proposition 1. *Let $A \in \mathbb{C}^{n\times n}$ be a matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, pairwise distinct. Let $v_1, v_2, \ldots, v_n$ and let $l_1, l_2, \ldots, l_n$ be their associated right and left eigenvectors, such that, $\|v_i\| = 1$, $i = 1, 2, \ldots, n$, and $l_i^* v_j = \delta_{ij}$. Let $s_1, s_2, \ldots, s_n$ be the corresponding eigenvalue condition numbers.*

Let $A^{(1)}$ be the matrix obtained by applying Theorem 1 to matrix A working with the right eigenvector associated with λ_1. Then the eigenvalue condition numbers of $A^{(1)}$ are given by

$$s_1^{(1)} = 1, \quad \text{and} \quad s_i^{(1)} = |\sin(\alpha_{1i})|\, s_i, \ \text{for } i = 2, 3, \ldots, n,$$

where α_{1i} denotes the angle between the vectors v_1 and v_i.

Proof. Let α_{1i} be the angle between the vectors v_1 and v_i, $i = 2, 3, \ldots, n$. By definition of eigenvalue condition number we have, for $i = 2, 3, \ldots, n$, that

$$s_i^{(1)} = \frac{\left\|v_i^{(1)}\right\| \left\|l_i^{(1)}\right\|}{\left|\left(l_i^{(1)}\right)^* v_i^{(1)}\right|} = \left\|v_i^{(1)}\right\| \|l_i\| = \left\|v_i^{(1)}\right\| s_i = |\sin(\alpha_{1i})|\, s_i.$$

□

Consequently, smaller angle between the vectors v_1 and v_i better eigenvalues condition number $s_i^{(1)}$ of $A^{(1)}$. Of course an alternative method can be used for instance choosing the eigenvalue with the largest condition number as we can done in Remark 1.

Next result gives an expression to compute $q_{(1)}$ by a matrix vector product. Note that this expression use only one eigenvalue and its right eigenvector.

Theorem 2. *The unique solution of the system (4) can be obtained directly by*

$$q_{(1)}^* = v_1^*(\lambda_1 I - A). \tag{7}$$

Proof. Consider the similar matrices A and $A^{(1)} = A + v_1 q_{(1)}^*$ of Theorem 1. Let $J_A = V^{-1}AV$ be the Jordan form of A, where

$$V = [v_1 \ v_2 \ \ldots \ v_n] \quad \text{and} \quad V^{-1} = \begin{bmatrix} l_1^* \\ l_2^* \\ \vdots \\ l_n^* \end{bmatrix},$$

with $\|v_i\| = 1, \ i = 1, 2, \ldots, n$. Then

$$J_A = J_{A^{(1)}} = \left(V^{(1)}\right)^{-1} A^{(1)} V^{(1)}.$$

By equations (5) and (6) we have

$$V^{(1)} = [v_1^{(1)} \ v_2^{(1)} \ \ldots \ v_n^{(1)}] = [v_1 \ v_2 - (v_1^* v_2) v_1 \ \ldots \ v_n - (v_1^* v_n) v_1],$$

$$\left(V^{(1)}\right)^{-1} = \begin{bmatrix} v_1^* \\ l_2^* \\ \vdots \\ l_n^* \end{bmatrix}.$$

Therefore, $A^{(1)} = \left(V^{(1)} V^{-1}\right) A \left(V \left(V^{(1)}\right)^{-1}\right) = T_1^{-1} A T_1$, where

$$\begin{aligned} T_1 &= [v_1 \ v_2 \ \ldots \ v_n] \begin{bmatrix} v_1^* \\ l_2^* \\ \vdots \\ l_n^* \end{bmatrix} = v_1 v_1^* + v_2 l_2^* + v_3 l_3^* + \cdots + v_n l_n^* \\ &= v_1 v_1^* + I - v_1 l_1^* = I + v_1 \left(v_1^* - l_1^*\right), \\ T_1^{-1} &= I - v_1 \left(v_1^* - l_1^*\right). \end{aligned}$$

Then,

$$\begin{aligned} A^{(1)} &= A + v_1 q_{(1)}^* = T_1^{-1} A T_1 = \left(I - v_1 \left(v_1^* - l_1^*\right)\right) A \left(I + v_1 \left(v_1^* - l_1^*\right)\right) \\ &= \left(I - v_1 \left(v_1^* - l_1^*\right)\right) \left(A + \lambda_1 v_1 \left(v_1^* - l_1^*\right)\right) \\ &= A + \lambda_1 v_1 \left(v_1^* - l_1^*\right) - v_1 v_1^* A + \lambda_1 v_1 l_1^* - \lambda_1 v_1 \left(v_1^* - l_1^*\right) v_1 \left(v_1^* - l_1^*\right) \\ &= A + \lambda_1 v_1 v_1^* - v_1 v_1^* A \\ &= A + v_1 \left(\lambda_1 v_1^* - v_1^* A\right). \end{aligned}$$

Note that, the vector $\lambda_1 v_1^* - v_1^* A$ satisfies

$$\left(\lambda_1 v_1^* - v_1^* A\right) v_1 = \lambda_1 v_1^* v_1 - v_1^* A v_1 = \lambda_1 \|v_1\|^2 - \lambda_1 \|v_1\|^2 = 0,$$

and for $i = 2, 3, \ldots, n$,

$$\left(\lambda_1 v_1^* - v_1^* A\right) v_i = \lambda_1 v_1^* v_i - v_1^* A v_i = \lambda_1 v_1^* v_i - \lambda_i v_1^* v_i = (\lambda_1 - \lambda_i)(v_1^* v_i).$$

Then, the system (4) has a unique solution

$$q_{(1)}^* = \lambda_1 v_1^* - v_1^* A = v_1^*(\lambda_1 I - A).$$

$\square$

Remark 2. Note that this rank one updated process can be applied recursively without losing the improved condition numbers. That is, with the matrix $A^{(1)} = A + v_1 q_{(1)}^*$ we obtain a rank one updated matrix $A^{(2)} = A^{(1)} + v_2^{(1)} q_{(2)}^*$, where $v_2^{(1)}$ is the right eigenvector of $A^{(1)}$ associated with λ_2 and where $q_{(2)}$ is obtained by the updated expression (7)

$$q_{(2)}^* = \left(v_2^{(1)}\right)^* \left(\lambda_2 I - A^{(1)}\right).$$

Now, the eigenvalue condition numbers of the eigenvalues of $A^{(2)}$, λ_1 and λ_2, are both equal to 1 and the remaining condition numbers are less than or equal to those of the initial matrix.

References

[1] Brauer, A.: Limits for the characteristic roots of matrices IV: Applications to stochastic matrices. Duke Math. J. 19, 75–91 (1952)

[2] Bru, R., Cantó, R., Soto, R.L., Urbano, A.M.: A Brauer's Theorem and related results. Central European Journal of Mathematics 10(1), 312–321 (2012)

[3] Bunchand, J.R., Nielsen, C.P.: Updating the singular value decomposition. Numerische Mathematik 31, 111–129 (1978)

[4] Bunchand, J.R., Nielsen, C.P., Sorensen, D.C.: Rank-one modification of the symmetric eigenproblem. Numerische Mathematik 31, 31–48 (1978)

[5] Byers, R., Kressner, D.: On the condition of a complex eigenvalue under real perturbations. BIT Numerical Mathematics 43(1), 1–18 (2003)

[6] Golub, G.H., Van Loan, C.F.: Matrix Computations. Johns Hopkins, Baltimore, fourth edition, (2013)

[7] Saad, Y.: Numerical Methods for large eigenvalue problems. SIAM, Philadelphia (2011)

[8] Watkins, D.S.: Fundamentals of Matrix Computations. John Wiley and Sons, New York (1991)

[9] Wilkinson, J.H.: The algebraic eigenvalue problem. Oxford University Press, (1965)