On Wilkinson's problem

A.N. Malyshev

Abstract

We prove that the 2-distance from an $n \times n$ matrix A to the matrices that have a multiple eigenvalue λ is equal to

$$\max_{\gamma \geq 0} \sigma_{2n-1} \begin{pmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{pmatrix},$$

where the singular values σ_k are ordered nonincreasingly.

1 Introduction

Given a matrix $A \in \mathbf{C^{n \times n}}$, $n \geq 2$, let us define the following parameter:

$$rsep(A) = \min \left\{ \|\Delta\|_2 \mid \Delta \in \mathbf{C^{n \times n}},\ A + \Delta \text{ has a multiple eigenvalue} \right\}. \tag{1}$$

Estimation of this parameter and its relation to the sensitivity of eigenvalues are much discussed in papers [5, 8, 9] and references therein.

The main result of the present paper is the derivation of the formula

$$rsep(A) = \min_{\lambda \in \mathbf{C}} \max_{\gamma \geq 0} \sigma_{2n-1} \begin{pmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{pmatrix}, \tag{2}$$

where I is the identity matrix and σ_k stands for the k-th singular value. The singular values are ordered nonincreasingly: $\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$.

As a matter of fact, formula (2) solves a more general problem than just the calculation of $rsep(A)$. Indeed, introduce the parameter

$$rsep_\lambda(A) = \max_{\gamma \geq 0} \sigma_{2n-1} \begin{pmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{pmatrix}, \tag{3}$$

then $rsep_\lambda(A)$ is equal to the distance from A to the matrices with a multiple eigenvalue λ. The contour lines of the function $p(\lambda) = rsep_\lambda(A)$ with the values $p(\lambda) = \epsilon \geq rsep(A)$ give us a "pseudospectrum of eigenvalue multiplicities".

In addition to formula (3) we give a constructive description of the matrix Δ of minimal 2-norm such that $A - \Delta$ has a multiple eigenvalue λ.

Similar results, inspired by the theory of aerodynamical flutter [6], are presented for a real matrix $A \in \mathbf{R^{n \times n}}$.

2 The distance to a multiple zero eigenvalue

It is evident that

$$rsep(A) = \min_{\lambda \in \mathbf{C}} rsep_\lambda(A), \qquad rsep_\lambda(A) = rsep_0(A - \lambda I),$$

where

$$rsep_0(A) = \min \left\{ \|\Delta\|_2 \mid \Delta \in \mathbf{C}^{\mathbf{n}\times\mathbf{n}},\ A + \Delta \text{ has a multiple zero eigenvalue} \right\}. \tag{4}$$

Therefore, in what follows we restrict ourselves to the problem of calculation of $rsep_0(A)$ for an arbitrary matrix $A \in \mathbf{C}^{\mathbf{n}\times\mathbf{n}}$.

By means of the Jordan canonical form of matrices one can prove the following

Lemma 1 *The matrix $A + \Delta$ has a multiple zero eigenvalue if and only if there exist linearly independent vectors $x \in \mathbf{C}$ and $y \in \mathbf{C}$ such that*

$$(A + \Delta)^2 x = 0 \qquad \textit{and} \qquad (A + \Delta)^2 y = 0.$$

Suppose now that $(A + \Delta)^2 x = 0$ and $(A + \Delta)^2 y = 0$ with some linearly independent vectors x and y. Let us introduce the vectors

$$x(\gamma) = -\frac{1}{\gamma}(A + \Delta)x = 0 \qquad \text{and} \qquad y(\gamma) = -\frac{1}{\gamma}(A + \Delta)y = 0$$

for any real $\gamma \neq 0$. Then

$$\begin{pmatrix} A + \Delta & \gamma I \\ 0 & A + \Delta \end{pmatrix} \begin{pmatrix} x \\ x(\gamma) \end{pmatrix} = 0 \qquad \text{and} \qquad \begin{pmatrix} A + \Delta & \gamma I \\ 0 & A + \Delta \end{pmatrix} \begin{pmatrix} y \\ y(\gamma) \end{pmatrix} = 0.$$

Since x and y are linearly independent, for any real $\gamma \neq 0$

$$\sigma_{2n-1} \begin{pmatrix} A + \Delta & \gamma I \\ 0 & A + \Delta \end{pmatrix} = 0.$$

Perturbation theory of singular values yields the inequality

$$\|\Delta\|_2 \geq \sigma_{2n-1} \begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix},$$

which holds for any real $\gamma \neq 0$. As a result, we derive the following estimate:

$$rsep_0(A) = \inf \|\Delta\|_2 \geq \sup_{\gamma \in \mathbf{R}} \sigma_{2n-1} \begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}. \tag{5}$$

Note that the value $\gamma \neq 0$ is included in (5) by continuity.

In order to prove the reverse inequality

$$rsep_0(A) = \min \|\Delta\|_2 \leq \max_{\gamma \in \mathbf{R}} \sigma_{2n-1} \begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix},$$

we explicitly construct a suitable perturbation matrix Δ with

$$\|\Delta\|_2 = \max_{\gamma\in\mathbf{R}} \sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}.$$

This approach was borrowed from [4].

Let us introduce for convenience the following notation:

$$F(\gamma) = \begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix} \qquad \text{and} \qquad f(\gamma) = \sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}.$$

3 Some properties of the function $f(\gamma)$

First of all, the function $f(\gamma) = \sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}$ is continuous and symmetric, i.e. $f(\gamma) = f(-\gamma)$.

Lemma 2

$$f(\gamma) \to 0 \qquad \text{as } |\gamma| \to \infty.$$

Proof: Assume first that A is nonsingular, then

$$\sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix} = 1/\sigma_2\begin{pmatrix} A^{-1} & -\gamma A^{-2} \\ 0 & A^{-1} \end{pmatrix}.$$

By the perturbation theory

$$\sigma_2\begin{pmatrix} A^{-1} & -\gamma A^{-2} \\ 0 & A^{-1} \end{pmatrix} \geq |\gamma|\sigma_2(A^{-2}) - \|A^{-1}\|,$$

therefore, $\sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix} \to 0$ as $|\gamma| \to \infty$.

Now assume that A is singular and fix any $\epsilon > 0$. Let us choose a nonsingular matrix A_ϵ such that $\|A_\epsilon - A\|_2 \leq \epsilon$. Owing to the above argument there exists $\gamma_0 > 0$ such that $\sigma_{2n-1}\begin{pmatrix} A_\epsilon & \gamma I \\ 0 & A_\epsilon \end{pmatrix} < \epsilon$, when $|\gamma| \geq \gamma_0$. Since $\sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix} \leq \sigma_{2n-1}\begin{pmatrix} A_\epsilon & \gamma I \\ 0 & A_\epsilon \end{pmatrix} + \epsilon$, we obtain the inequality $\sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix} < 2\epsilon$, when $|\gamma| \geq \gamma_0$. $\square$

The next argument shows that either $f(\gamma) \neq 0$ for all $\gamma \neq 0$ or $f(\gamma) \equiv 0$. Indeed, suppose that $\sigma_{2n-1}\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix} = 0$ for some real $\gamma \neq 0$. Then there exist linearly independent vectors $\begin{pmatrix} x \\ x_1 \end{pmatrix}$ and $\begin{pmatrix} x \\ x_1 \end{pmatrix}$ such that

$$\begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}\begin{pmatrix} x \\ x_1 \end{pmatrix} = 0 \qquad \text{and} \qquad \begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}\begin{pmatrix} y \\ y_1 \end{pmatrix} = 0.$$

These vectors are right singular vectors corresponding to $\sigma_{2n-1} = \sigma_{2n} = 0$. Trivial calculations show that $A^2 x = 0$ and $A^2 y = 0$. Now we demonstrate that the vectors x and y are linearly independent. Suppose from the contrary that there exist scalars α and β, at least one of them being non-zero, such that $\alpha x + \beta y = 0$. Then, obviously, $\alpha x_1 + \beta y_1 = -\frac{1}{\gamma} A(\alpha x + \beta y) = 0$ and $\alpha \begin{pmatrix} x \\ x_1 \end{pmatrix} + \beta \begin{pmatrix} y \\ y_1 \end{pmatrix} = 0$, which leads to a contradiction. Thus, by Lemma 1 the matrix A has a multiple zero eigenvalue and $f(\gamma) \equiv 0$.

4 Properties of singular vectors of the matrix $F(\gamma)$ at a local extremum

Assume that $\gamma_* \in \mathbf{R}$ is a local extremum of $f(\gamma)$ and

$$\sigma_{2n-1} \begin{pmatrix} A & \gamma_* I \\ 0 & A \end{pmatrix} = \sigma_* > 0.$$

There exist vectors $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ of unit length such that

$$\begin{pmatrix} A & \gamma_* I \\ 0 & A \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \sigma_* \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \tag{6}$$

$$\begin{pmatrix} A^H & 0 \\ \gamma_* I & A^H \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \sigma_* \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{7}$$

where A^H denotes the Hermitian conjugate of A.

The vectors u_1, u_2, v_1, and v_2 possess a number of remarkable properties in addition to $u_1^H u_1 + u_2^H u_2 = 1 = v_1^H v_1 + v_2^H v_2$.

Lemma 3 *If (6) and (7) hold with $\sigma_* > 0$, then*

$$u_1^H u_2 = v_1^H v_2. \tag{8}$$

Proof: Multiplying (6) from the left by $[0\; v_1^H]$ and (7) by $[u_2^H\; 0]$, we obtain $v_1^H A u_2 = \sigma_* v_1^H v_2$ and $u_2^H A^H v_1 = \sigma_* u_2^H u_1$. As a result, $\sigma_*(u_1^H u_2 - v_1^H v_2) = 0$. □

The following lemma is a corollary of classical results [2] and can be proved as in [1].

Lemma 4 *Let $F(\gamma) \in \mathbf{C}^{\mathbf{m}\times\mathbf{n}}$ be an analytic matrix function on an open set $\Gamma \subset \mathbf{R}$. Then there exist an analytic diagonal matrix function $\tilde{\Sigma}(\gamma) = diag(\tilde{\sigma}_1(\gamma), \ldots, \tilde{\sigma}_{\min\{m,n\}}(\gamma)) \in \mathbf{R}^{\mathbf{m}\times\mathbf{n}}$ and analytic unitary matrix functions $\tilde{U}(\gamma) = [\tilde{u}_1(\gamma), \ldots, \tilde{u}_n(\gamma)] \in \mathbf{C}^{\mathbf{n}\times\mathbf{n}}$ and $\tilde{V}(\gamma) = [\tilde{v}_1(\gamma), \ldots, \tilde{v}_m(\gamma)] \in \mathbf{C}^{\mathbf{m}\times\mathbf{m}}$, all of which are defined on Γ, such that $\tilde{\Sigma}(\gamma) = \tilde{V}^H(\gamma) F(\gamma) \tilde{U}(\gamma)$.*

It follows from Lemma 4 [7] that

$$\frac{d\tilde{\sigma}_i(\gamma)}{d\gamma} = \Re\left(\tilde{v}_i^H(\gamma) \frac{dF(\gamma)}{d\gamma} \tilde{u}_i(\gamma) \right), \tag{9}$$

for $i = 1, \ldots, \min\{m,n\}$. Indeed, differentiating the identity $\tilde{\sigma}_i(\gamma) = \tilde{v}_i^H(\gamma)F(\gamma)\tilde{u}_i(\gamma)$ with respect to γ yields

$$\frac{d\tilde{\sigma}_i(\gamma)}{d\gamma} = \tilde{v}_i^H(\gamma)\frac{dF(\gamma)}{d\gamma}\tilde{u}_i(\gamma) + \tilde{\sigma}_i(\gamma)\left[\frac{\tilde{v}_i^H(\gamma)}{d\gamma}\tilde{v}_i(\gamma) + \tilde{u}_i^H(\gamma)\frac{\tilde{u}_i(\gamma)}{d\gamma}\right].$$

Since $\tilde{u}_i^H(\gamma)\tilde{u}_i(\gamma) = \tilde{v}_i^H(\gamma)\tilde{v}_i(\gamma) = 1$, we derive that $\Re\left(\tilde{u}_i^H(\gamma)\frac{d\tilde{u}_i(\gamma)}{d\gamma}\right) = \Re\left(\frac{d\tilde{v}_i^H(\gamma)}{d\gamma}\tilde{v}_i(\gamma)\right) = 0$.

It is evident that $|\tilde{\sigma}_i(\gamma)|$, $i = 1, \ldots, \min\{m,n\}$, are the singular values of $F(\gamma)$. Recall that the notation $\sigma_1(\gamma) \geq \ldots \geq \sigma_{\min\{m,n\}}(\gamma) \geq 0$ is reserved for the nonincreasingly ordered singular values, which are continuous and piecewise-analytic functions on Γ.

The formulation and proof of the following lemma is taken from [4].

Lemma 5 *Let $F(\gamma) \in \mathbf{C}^{\mathbf{m}\times\mathbf{n}}$ be an analytic matrix function on an open set $\Gamma \subset \mathbf{R}$. Let $\sigma_1(\gamma) \geq \ldots \geq \sigma_{\min\{m,n\}}(\gamma) \geq 0$ be its ordered singular values. If $\sigma_i(\gamma_*) > 0$ at a local extremum $\gamma_* \in \Gamma$, then there exists a pair of left and right singular vectors $v \in \mathbf{C}^{\mathbf{m}}$ and $u \in \mathbf{C}^{\mathbf{n}}$ of $F(\gamma_*)$ corresponding to $\sigma_i(\gamma_*)$ such that $\Re\left(v^H\frac{dF}{d\gamma}(\gamma_*)u\right) = 0$.*

Proof: If γ_* is an analyticity point of $\sigma_i(\gamma)$, then $\sigma_i(\gamma) = \tilde{\sigma}_i(\gamma)$ in a neighbourhood of γ_*, and by Lemma 4 $\Re\left(\tilde{v}_i^H(\gamma_*)\frac{dF}{d\gamma}(\gamma_*)\tilde{u}_i(\gamma_*)\right) = 0$.

Otherwise, the function $\sigma_i(\gamma)$ can be represented in a neighbourhood of γ_* as $\sigma_i(\gamma) = \tilde{\sigma}_l(\gamma)$ if $\gamma \leq \gamma_*$ and $\sigma_i(\gamma) = \tilde{\sigma}_r(\gamma)$ if $\gamma \geq \gamma_*$ so that either $\frac{d\tilde{\sigma}_l}{d\gamma}(\gamma_*) \leq 0$ and $\frac{d\tilde{\sigma}_r}{d\gamma}(\gamma_*) \geq 0$ or $\frac{d\tilde{\sigma}_l}{d\gamma}(\gamma_*) \geq 0$ and $\frac{d\tilde{\sigma}_r}{d\gamma}(\gamma_*) \leq 0$.

Let us search for the vectors u and v in the form $u_\alpha = \alpha\tilde{u}_l(\gamma_*) + \sqrt{1-\alpha^2}\tilde{u}_r(\gamma_*)$, $v_\alpha = \alpha\tilde{v}_l(\gamma_*) + \sqrt{1-\alpha^2}\tilde{v}_r(\gamma_*)$, where $\alpha \in [0,1]$. Since $\frac{d\tilde{\sigma}_l}{d\gamma}(\gamma_*) = \Re\left(v_1^H\frac{dF}{d\gamma}(\gamma_*)u_1\right)$ and $\frac{d\tilde{\sigma}_r}{d\gamma}(\gamma_*) = \Re\left(v_0^H\frac{dF}{d\gamma}(\gamma_*)u_0\right)$, the continuous function $\phi(\alpha) = Re\left(v_\alpha^H\frac{dF}{d\gamma}(\gamma_*)u_\alpha\right)$ satisfies $\phi(1)\phi(0) \leq 0$. Therefore, there exists $\alpha_* \in [0,1]$ such that $\phi(\alpha_*) = 0$. $\square$

Let us apply Lemma 5 to the matrix $F(\gamma) = \begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}$. As a result, there exists a pair of singular vectors $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ of $F(\gamma_*)$ such that

$$0 = \Re\left((v_1^H\ v_2^H)\frac{dF}{d\gamma}(\gamma_*)\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = \Re\left((v_1^H\ v_2^H)\begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = \Re\left(v_1^H u_2\right).$$

Multiplying (6) from the left by $[v_1^H \ \ -v_2^H]$ and (7) by $[u_1^H \ \ -u_2^H]$, we obtain the identities

$$v_1^H A u_1 - v_2^H A u_2 + \gamma_* v_1^H u_2 = \sigma_*(v_1^H v_1 - v_2^H v_2),$$
$$u_1^H A^H v_1 - u_2^H A^H v_2 - \gamma_* u_2^H v_1 = \sigma_*(u_1^H u_1 - u_2^H u_2).$$

Comparing them, we derive

$$\sigma_*(v_1^H v_1 - v_2^H v_2) = \sigma_*(u_1^H u_1 - u_2^H u_2) + 2\gamma_* v_1^H u_2.$$

Hence $\sigma_*(v_1^H v_1 - v_2^H v_2) = \sigma_*(u_1^H u_1 - u_2^H u_2)$ and $\gamma_* v_1^H u_2 = 0$. Taking into account that $\sigma_* > 0$ and $v_1^H v_1 + v_2^H v_2 = 1 = u_1^H u_1 + u_2^H u_2$, we get

$$v_1^H v_1 = v_2^H v_2 \qquad \text{and} \qquad u_1^H u_1 = u_2^H u_2.$$

All the results of this section can be summarized in

Lemma 6 *If γ_* is a local extremum of $f(\gamma)$ and $f(\gamma_*) = \sigma_* > 0$, then there exist singular vectors $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ of unit length for the matrix $F(\gamma_*)$ such that $\Re\left(v_1^H u_2\right) = 0$. Furthermore, the matrices $U = [u_1 u_2]$ and $V = [v_1 v_2]$, composed of these vectors, satisfy $U^H U = V^H V$.*

5 Construction of a minimal perturbation Δ

5.1 $\gamma_* \neq 0$

Let us use the matrices $U = [u_1 u_2]$ and $V = [v_1 v_2]$ from Lemma 6. One can show that the vectors u_1 and u_2 are linearly independent. From the identity $U^T U = V^T V$ it follows that $\alpha u_1 + \beta u_2 = 0$ implies $\alpha v_1 + \beta v_2 = 0$ and vice versa. Suppose, for instance, that $\alpha \neq 0$ and $\alpha u_1 + \beta u_2 = 0$, then $u_1 = -\frac{\beta}{\alpha} u_2$, $v_1 = -\frac{\beta}{\alpha} v_2$. From (6) we derive that $A u_1 + \gamma_* u_2 = \sigma_* v_1$ and $A u_2 = \sigma_* v_2$. Therefore, $\gamma_* u_2 = 0$, $u_2 = 0$, and $u_1 = 0$. This contradicts with $u_1^H u_1 + u_2^H u_2 = 1$. Similarly, if $\beta \neq 0$, then $u_2 = -\frac{\alpha}{\beta} u_1$, $v_2 = -\frac{\alpha}{\beta} v_1$. It follows from (7) that $A^H v_1 = \sigma_* u_1$ and $\gamma_* v_1 + A^H v_2 = \sigma_* u_2$. Therefore, $\gamma_* v_1 = 0$, $v_1 = 0$, and $v_2 = 0$. Thus we come at contradiction with $v_1^H v_1 + v_2^H v_2 = 1$.

Let us construct Δ in the form $\Delta = -\sigma_* V U^+$, where U^+ is the Moore-Penrose pseudoinverse of U. Since $V^T V = U^T U$, there exists a unitary matrix W such that $V = WU$. Therefore, $\|\Delta\|_2 = \sigma_* \|V U^+\|_2 = \sigma_* \|U U^+\|_2 = \sigma_*$ because $U U^+$ is a unitary projector.

Furthermore, $(A + \Delta)U = [-\gamma_* u_2,\ 0]$ and, consequently, $(A + \Delta)^2 U = 0$. By Lemma 1 $A + \Delta$ has a multiple zero eigenvalue.

5.2 $\gamma_* = 0$

The condition $\sigma_{2n-1}\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} = \sigma_* > 0$ implies that $\sigma_n(A) = \sigma_* > 0$. There are two cases: $\sigma_{n-1}(A) = \sigma_n(A)$ or $\sigma_{n-1}(A) > \sigma_n(A)$.

5.2.1 Case 1: $\sigma_{n-1}(A) = \sigma_n(A)$

If $\sigma_{n-1}(A) = \sigma_n(A) = \sigma_* > 0$, then there exist right and left singular vectors u_1, u_2, v_1, v_2 of A corresponding to these eigenvalues such that $U^H U = I = V^H V$, where $U = [u_1 u_2]$, $V = [v_1 v_2]$.

The matrix $\Delta = -\sigma_* V U^H$ possesses the following properties: $\|\Delta\|_2 = \sigma_*$, $(A + \Delta)U = 0$. Therefore, the matrix $A + \Delta$ has a multiple zero eigenvalue.

5.2.2 Case 2: $\sigma_{n-1}(A) > \sigma_n(A)$

The situation, when $\sigma_{n-1}(A) > \sigma_n(A) = \sigma_* > 0$, is rather involved. Denote by u, v a pair of singular vectors of unit length such that $Au = \sigma_* v$, $A^T v = \sigma_* u$. Let us prove that $v^H u = 0$. Only here we use the fact that the local extremum at $\gamma_* = 0$ is a local maximum.

Consider the analytic matrix function $F(\gamma) = \begin{pmatrix} A & \gamma I \\ 0 & A \end{pmatrix}$ for such small $|\gamma|$, where $\sigma_{2n-2}(F(\gamma)) > \sigma_{2n-1}(F(\gamma))$. This range of γ we denote by Γ. The matrix $F(\gamma)$ has analytic singular values $\tilde{\sigma}_{2n-1}(\gamma)$ and $\tilde{\sigma}_{2n}(\gamma)$ satisfying $\tilde{\sigma}_{2n-1}(0) = \tilde{\sigma}_{2n}(0) = \sigma_*$. Diminish Γ if needed so that $\tilde{\sigma}_{2n-1}(\gamma) \le \sigma_*$ and $\tilde{\sigma}_{2n}(\gamma) \le \sigma_*$ in the neighbourhood $\Gamma \subset \mathbf{R}$. Denote by $\tilde{u}_{2n-1}(\gamma)$, $\tilde{v}_{2n-1}(\gamma)$, $\tilde{u}_{2n}(\gamma)$, $\tilde{v}_{2n}(\gamma)$ the unitary analytic singular vectors of $F(\gamma)$ corresponding to $\tilde{\sigma}_{2n-1}(\gamma)$ and $\tilde{\sigma}_{2n}(\gamma)$.

Since $\tilde{\sigma}_{2n-1}(0)$ and $\tilde{\sigma}_{2n}(0) = \sigma_* = \sigma_n(A) < \sigma_{n-1}(A)$, one can prove that

$$\tilde{u}_{2n-1}(0) = \begin{pmatrix} k_{2n-1} u \\ l_{2n-1} u \end{pmatrix}, \quad \tilde{u}_{2n}(0) = \begin{pmatrix} k_{2n-1} u \\ l_{2n} u \end{pmatrix},$$

where the matrix $\begin{pmatrix} k_{2n-1} & k_{2n} \\ l_{2n-1} & l_{2n} \end{pmatrix}$ is unitary. Choose the scalars α and β that satisfy the system $\begin{pmatrix} k_{2n-1} & k_{2n} \\ l_{2n-1} & l_{2n} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ u^H v \end{pmatrix} \frac{1}{\sqrt{1+|v^H u|^2}}$. Therefore, $\alpha^H \alpha + \beta^H \beta = 1$. Denote $x(\gamma) = \alpha \tilde{u}_{2n-1}(\gamma) + \beta \tilde{u}_{2n}(\gamma)$ and $y(\gamma) = \alpha \tilde{v}_{2n-1}(\gamma) + \beta \tilde{v}_{2n}(\gamma)$. Then $x^H(\gamma)x(\gamma) = 1$, $F(\gamma)x(\gamma) = [\tilde{v}_{2n-1}(\gamma), \tilde{v}_{2n}(\gamma)] \begin{pmatrix} \alpha \tilde{\sigma}_{2n-1}(\gamma) \\ \beta \tilde{\sigma}_{2n}(\gamma) \end{pmatrix}$, $\|F(\gamma)x(\gamma)\|_2 \le \sigma_*$, $F(0)x(0) = \sigma_* y(0)$, $F^H(0)y(0) = \sigma_* x(0)$, $\|F(0)x(0)\| = \sigma_*$. Differentiate $x^T(\gamma)F^T(\gamma)F(\gamma)x(\gamma)$ at $\gamma = 0$:

$$\begin{aligned} \frac{d}{d\gamma}\left(x^T(\gamma)F^T(\gamma)F(\gamma)x(\gamma)\right)(0) &= \sigma_*^2 \left(\frac{dx^H}{d\gamma}(0)x(0) + x^H(0)\frac{dx^H}{d\gamma}(0)\right) + \\ & \sigma_* \left[x^H(0) \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix} y(0) + y^H(0) \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} x(0)\right]. \end{aligned}$$

The identity $x^H(\gamma)x(\gamma) = 1$ implies $\frac{dx^H}{d\gamma}(0)x(0) + x^H(0)\frac{dx^H}{d\gamma}(0) = 0$. Since

$$x(0) = \frac{1}{\sqrt{1+|v^H u|^2}} \begin{pmatrix} u \\ (u^H v)u \end{pmatrix} \quad \text{and} \quad y(0) = \frac{1}{\sqrt{1+|v^H u|^2}} \begin{pmatrix} v \\ (u^H v)v \end{pmatrix},$$

we obtain the equality $y^H(0) \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} x(0) = \frac{|v^H u|^2}{1+|v^H u|^2} = 0$. As a result,

$$\frac{d}{d\gamma}\left(x^T(\gamma)F^T(\gamma)F(\gamma)x(\gamma)\right)(0) = \sigma_* \frac{2|v^H u|^2}{1+|v^H u|^2},$$

which yields $v^H u = 0$.

Let us choose now $\Delta = -\sigma_* v u^H$. It follows immediately that $(A+\Delta)u = 0$. A necessary and sufficient condition of solvability of the equation $(A+\Delta)w = u$ with respect to w is the identity $(A+\Delta)(A+\Delta)^+u = u$. It is not hard to verify that $(A+\Delta)(A+\Delta)^+ = I - vv^H$, therefore, $(A+\Delta)(A+\Delta)^+u = (I-vv^H)u = u$. The vectors u and w are linearly independent. Indeed, if $\alpha w + \beta u = 0$, then multiplying it from the left by $A+\Delta$ we obtain $\alpha u = 0$. Since $u \neq 0$, $\alpha = 0$. Furthermore, $\beta u = 0$ and $\beta = 0$. Finally, apply Lemma 1 because $(A+\Delta)^2[u\, w] = 0$.

The case $\gamma_* = 0$ with $\sigma_{n-1}(A) > \sigma_n(A)$ appears to be generic when searching for $\min_\lambda rsep_\lambda(A)$.

5.3 Extremal properties of $f(\gamma)$

Assume that $f(\gamma) \not\equiv 0$. As shown above, for every local extremum $\gamma_* \neq 0$ of the function $f(\gamma)$ one can construct a suitable perturbation Δ such that $A+\Delta$ has a multiple zero eigenvalue and $\|\Delta\|_2 = f(\gamma_*)$. Therefore, $f(\gamma_*) = \|\Delta\|_2 \geq \max_{\gamma\in\mathbf{R}} f(\gamma)$, i.e. γ_* is necessarily a global maximum of $f(\gamma)$. Thus, either $f(\gamma)$ has no local extrema on $(0,\infty)$ or it has only one, which is a global maximum. In the former case the global maximum of $f(\gamma)$ on $[0,\infty)$ is attained at $\gamma = 0$.

The function $f(\gamma)$ possesses two typical plot patterns:

Behaviour of the function $rsep_\lambda(A)$ with respect to λ can be rather involved (compare, e.g., with the standard pseudospectrum). Therefore, computation of its extremal points such as $rsep(A)$ might be challenging.

6 On stability of the spectral structure of real matrices

The flutter analysis [6] often applies the following problem of linear algebra. Given a real matrix $A \in \mathbf{R}^{\mathbf{n}\times\mathbf{n}}$, find out whether all eigenvalues of A are real. Our goal here is to introduce a measure of stability of such spectral property.

Assume that some matrix $A \in \mathbf{R}^{\mathbf{n}\times\mathbf{n}}$ has only real eigenvalues. Adding real perturbations Δ of continuously increasing norm to A, one can get matrices $A+\Delta$ that have complex eigenvalues. The break of the structure appears first, when some of the eigenvalues of $A+\Delta$ coalesce on the real axis. This argument allows us to define the parameter

$$d_{crit}(A) = \min\{\|\Delta\|_2 |\ A+\Delta \text{ has a multiple real eigenvalue}\}. \tag{10}$$

Repeating all proofs for the complex case and taking into account that one can always choose real singular vectors for real matrices, we derive the formula

$$d_{crit} = \min_{\lambda \in \mathbf{R}} \max_{\gamma \geq 0} \sigma_{2n-1} \begin{pmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{pmatrix}. \tag{11}$$

There exists a slightly generalized spectral problem, which concerns stable eigendecompositions of real matrices. Namely, it is easy to prove that the spectrum of a real matrix A is symmetric with respect to the real axis. Therefore, it can be divided into three parts: all eigenvalues above the real axis, all eigenvalues on the axis, and all eigenvalues below the axis. If we continuously change elements of A, a possible break of the spectral structure (the number of eigenvalues in each part) can appear only at the moment when two eigenvalues coalesce on the real axis. All the above results are applicable to this problem too, and the measure of stability of the spectral structure is equal to $\min_{\lambda \in \mathbf{R}} \max_{\gamma \geq 0} \sigma_{2n-1} \begin{pmatrix} A - \lambda I & \gamma I \\ 0 & A - \lambda I \end{pmatrix}$.

Concerning numerical computation of d_{crit}, it is more convenient [3] to use the formula

$$d_{crit} = \min_{\lambda \in \mathbf{R}} \{\sigma \mid \sigma = \sigma_{\min}(A - \lambda I), v^T u = 0, \text{ where } (A - \lambda I)u = \sigma v \text{ and } (A - \lambda I)^T v = \sigma u\}.$$

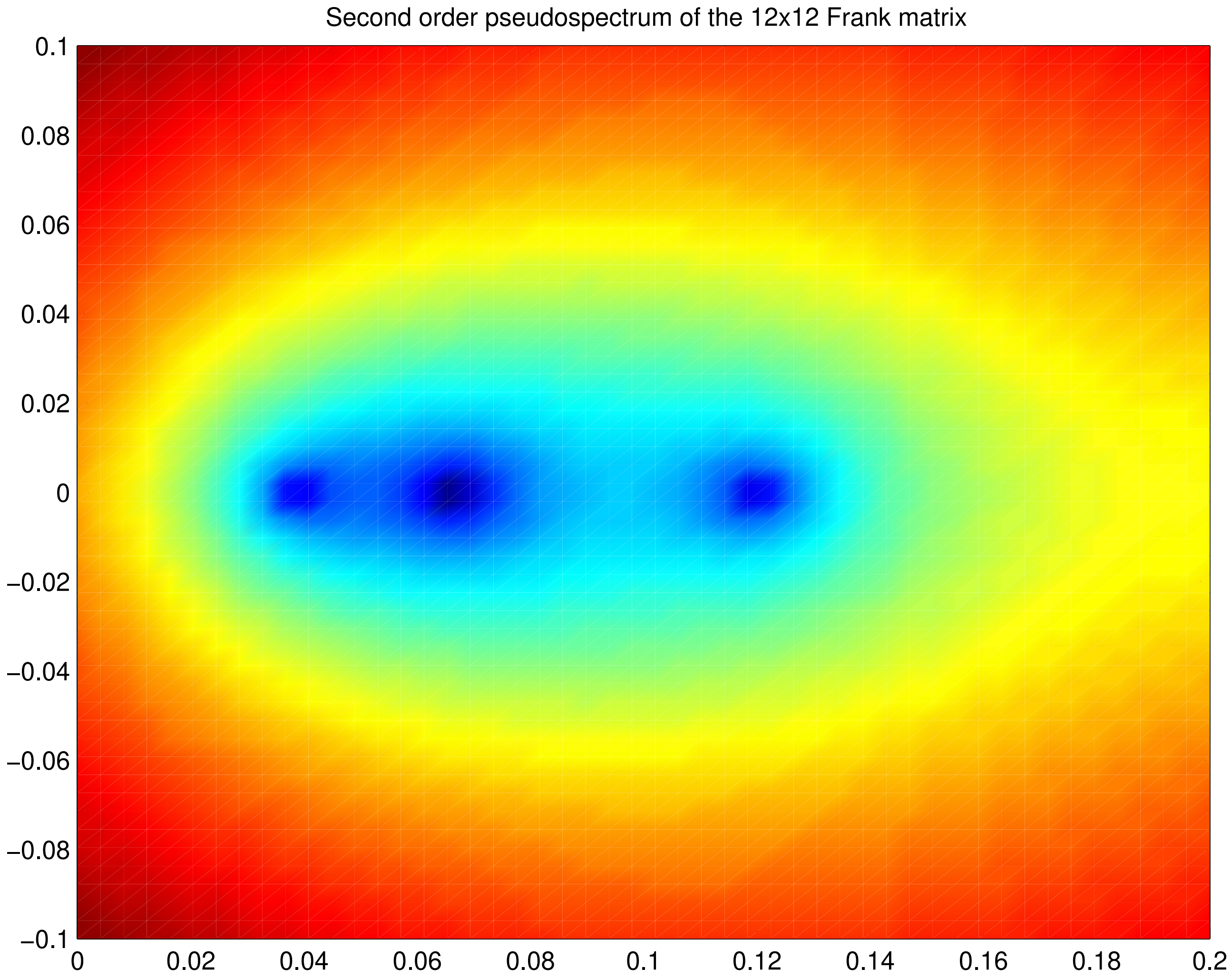
7 Numerical example

Consider a 12×12 Frank matrix from MATLAB's gallery:

$$A = \begin{pmatrix} 12 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 11 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 10 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 9 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 8 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 7 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 6 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

The eigenvalues of A are 0.0310, 0.0495, 0.0812, 0.1436, 0.2847, 0.6435, 1.5540, 3.5119, 6.9615, 12.3111, 20.1990, 32.2289. The corresponding reciprocals of condition numbers of the eigenvalues are $5.469e-8$, $2.579e-8$, $3.753e-8$, $1.492e-7$, $1.785e-6$, $6.912e-5$, 0.0046, 0.1445, 0.5845, 0.3182, 0.2008, 0.3042.

We computed a portrait of $rsep_\lambda(A)$ on the rectangle $0 \leq \Re(\lambda) \leq 0.2$, $-0.1 \leq \Im(\lambda) \leq 0.1$ with the uniform 50×50 mesh:



Acknowledgements. The author would like to thank Alan Edelman and Ross Lippert for many enlightening discussions of this paper.

References

[1] A. Bunse-Gerstner, R. Byers, V. Merhmann, and N.K. Nichols, Numerical computation of an analytic singular value decomposition of a matrix valued function, *Numer.Math.*, **60**, 1–39, 1991.

[2] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, 1980.

[3] R. Lippert, Notes on geometry of eigenvalues, 1997.

[4] Li Qui, B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young and J.C. Doyle, A formula for computation of the real stability radius, *Automatica*, **31**(6), 879–890, 1995.

[5] A. Ruhe, Properties of a matrix with a very ill-conditioned eigenproblem, *Numer.Math.*, **15**, 57–60, 1970.

[6] G. Sander, C. Bon and M. Geradin, Finite element analysis of supersonic panel flutter, *Int.J.Num.Meth.Engng.*, **7**, 379–394, 1973.

[7] Ji-guang Sun, A note on simple non-zero singular values, *J.Comput.Math.*, **6**, 259–266, 1988.

[8] J.H. Wilkinson, Note on matrices with a very ill-conditioned eigenproblem, *Numer.Math.*, **19**, 175–178, 1972.

[9] J.H. Wilkinson, On neighbouring matrices with quadratic elementary divisors, *Numer.Math.*, **44**, 1–21, 1984.