ON THE FIELD OF VALUES AND PSEUDOSPECTRA OF THE INVERSE OF A LARGE MATRIX

MICHIEL E. HOCHSTENBACH*, DAVID A. SINGER†, AND PAUL F. ZACHLIN‡

Abstract. We consider the approximation of the field of values of the inverse of a large sparse matrix, without explicitly computing the inverse or using its action (i.e., accurately solving a linear system with this matrix). We review results by Manteuffel and Starke and give an alternative that may be better approximations in practice. We give connections with the harmonic Rayleigh–Ritz approach and also study the approximation of the pseudospectra of a matrix inverse. We give several applications of the studied concepts as well as numerical examples.

Key words. Field of values, pseudospectra, matrix inverse, large sparse matrix, Ritz values, harmonic Rayleigh–Ritz, harmonic Ritz values, GMRES convergence, Arnoldi, numerical radius, numerical abscissa, inclusion region, exclusion region, matrix powers, matrix functions.

AMS subject classifications. 65F15, 65F50, 47A12, 65F30, 53A35, 30D50, 65F10.

1. Introduction.

If we would like to numerically approximate the field of values

$$\mathrm{FOV}(A) = \{x^*Ax : \|x\| = 1\},$$

where A is an $n \times n$ matrix and $\|\cdot\|$ denotes the two-norm, we may compute

$$\tfrac{1}{2}\ \lambda_{\max}(e^{i\alpha}A + (e^{i\alpha}A)^*)$$

for a number of angles α as initially proposed by Johnson [6]. Hereby the equality

$$\max_{z \in \mathrm{FOV}(A)} \mathrm{Re}(z) = \tfrac{1}{2}\,\lambda_{\max}(A + A^*) \tag{1.1}$$

is used for every angle.

In this paper we are interested in large sparse $n \times n$ matrices A. In this case $e^{i\alpha}A + (e^{i\alpha}A)^*$ is large, sparse, and Hermitian, and we may use the Lanczos method (see [7] and, e.g., [16]) to approximate the largest eigenvalue. This method generates a low-dimensional subspace to approximate eigenpairs of possibly large matrices. In principle we can run a new Lanczos process for every α in a chosen set; for each angle the largest eigenpair will be approximated using a different Krylov subspace, generated by a different matrix of the form $e^{i\alpha}A + (e^{i\alpha}A)^*$ and an initial vector, for instance a random vector or the approximate eigenvector for a previous value of α. We note hereby that since the Lanczos algorithm generally approximates the extremal eigenvalues well, we can approximate the maximal eigenvalues for the angles α and $\alpha+\pi$ simultaneously, realizing that $\lambda_{\max}(e^{i(\alpha+\pi)}A+(e^{i(\alpha+\pi)}A)^*) = -\lambda_{\min}(e^{i\alpha}A+(e^{i\alpha}A)^*)$.

An alternative, computationally less expensive, approach is to use only *one* subspace for all angles α as follows. We perform a single run of the Arnoldi process (see [1] and, e.g., [16]) on A (or on $e^{i\alpha}A$ for any fixed angle α) and an initial vector u_1 of unit length (for instance a random vector). This gives the Arnoldi decomposition

$$AU_k = U_k H_k + h_{k+1,k} u_{k+1} e_k^* = U_{k+1} \underline{H}_k, \tag{1.2}$$

*Department of Mathematics and Computing Science, Eindhoven University of Technology, PO Box 513, 5600 MB, The Netherlands. m.e.hochstenbach @tue. nl. Part of this work was done when the author was at Case Western Reserve University. The research of this author was supported in part by NSF grant DMS-0405387.

†Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106-7058, USA david.singer@case.edu

‡Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106-7058, USA paul.zachlin@case.edu

where the columns of U_k form an orthonormal basis for $\mathcal{U}_k$ with u_1 as its first column, H_k is an upper Hessenberg matrix, e_k is the kth canonical basis vector, and $\underline{H}_k = \begin{bmatrix} H_k \\ h_{k+1,k} e_k^* \end{bmatrix}$ is a $(k+1) \times k$ Hessenberg matrix with an extra row.

As originally suggested by Manteuffel and Starke [8] one can approximate $\mathrm{FOV}(A)$ by $\mathrm{FOV}(H_k)$, whereby we have

$$\mathrm{FOV}(A) \supseteq \mathrm{FOV}(U_k^* A U_k) = \mathrm{FOV}(H_k).$$

In particular, we know after k steps that $\mathrm{FOV}(A)$ contains the convex hull of the Ritz values, which are the eigenvalues of H_k. Since k is assumed to be much smaller than n (which is the main idea of a subspace method), determining $\mathrm{FOV}(H_k)$ is computationally very easy, for instance by the method proposed by Johnson.

We note that this second approach, working with H_k, is in practice often very sensible. While a different search space per α may often give a (slightly) better result (that is, a larger maximal eigenvalue per angle and hence a larger approximate FOV which is still included in the true FOV), both approaches give exactly or nearly the same approximation in the following examples.

EXAMPLE 1.1. Let A be an $n \times n$ tridiagonal matrix, and choose the initial vector $b = e_1$ (the first canonical basis vector) for the Krylov spaces $\mathcal{K}_k\left(\frac{1}{2}(e^{i\alpha}A + e^{-i\alpha}A^*), b\right)$ for each α ("first approach"), and for the Krylov space $\mathcal{K}_k(A, b)$ ("second approach"). Then it can be checked that in general (that is, apart from some exceptional matrices such as the zero matrix) all Krylov spaces are equal to $\mathrm{span}\{e_1, \ldots, e_k\}$ and, as a result, the approximations to the field of values will be exactly the same in both approaches. ⊘

EXAMPLE 1.2. As a second example, consider the 256×256 `grcar` matrix [9], which is banded upper Hessenberg; this example is also used in [8]. The eigenvalues and FOV are depicted in Figure 1.1(a). From Figure 1.1(b), we see that taking different Krylov subspaces $\mathcal{K}_k\left(\frac{1}{2}(e^{i\alpha}A + e^{-i\alpha}A^*), b\right)$ for each of the 32 angles ("first approach") instead of one Krylov subspace $\mathcal{K}_k(A, b)$ for all angles ("second approach") hardly improves the numerical approximation to the FOV, while the first approach is much more expensive. Here the initial vector b is random. ⊘

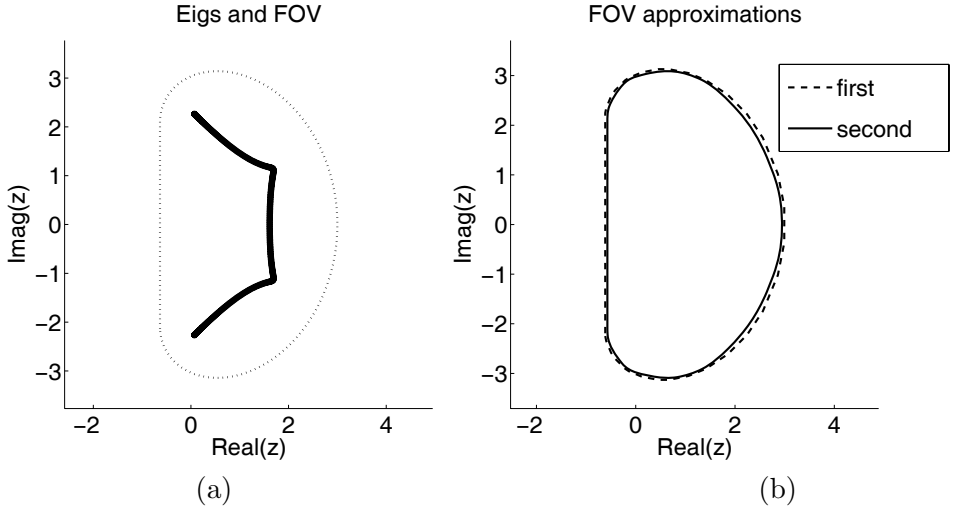


FIG. 1.1. *(a) Eigenvalues and FOV of the 256×256 `grcar` matrix. (b) Approximations of the FOV using different Krylov subspaces for each of the 32 angles ("first approach"), respectively one Krylov subspace for all angles ("second approach").*

In the remainder of this paper we concentrate on the field of values and pseudospectra for the inverse of a large sparse invertible matrix A. Hereby, we do not want to compute A^{-1}

explicitly, since this is often prohibitively expensive, but we also want to avoid the costs of accurately solving a linear system of the form $Ax = b$, as well as computing the maximal eigenvalues for a family of matrices, depending on an angle α as above.

In Section 2 we first give some applications of the field of values of a matrix inverse. We review an approach of Manteuffel and Starke and give relations with the harmonic Rayleigh–Ritz approach. We also present an alternative approach that comes with no inclusion guarantee, but may work better in practice. In addition, we provide some analysis, extensions, and several numerical examples.

Section 3 considers the pseudospectra of a matrix inverse: applications and four different techniques, that are compared analytically and in a numerical experiment. We end with some concluding remarks in Section 4.

2. The field of values of the inverse of a matrix. Suppose that we would like to approximate the field of values of A^{-1} without explicitly computing A^{-1}, or implicitly using its action (that is, exactly solving a system of the form $Ax = b$). We first mention a few applications where estimates for the field of values of A^{-1} may be useful.

2.1. Applications. First, we recall the definitions (see, e.g., [15]) of the *inner numerical radius*

$$\nu(A) = \min_{z \in \mathrm{FOV}(A)} |z|,$$

the *numerical radius*

$$\mu(A) = \max_{z \in \mathrm{FOV}(A)} |z|,$$

and the the *numerical abscissa*

$$\omega(A) = \max_{z \in \mathrm{FOV}(A)} \mathrm{Re}(z).$$

From (1.1) we know that $\omega(A) = \frac{1}{2}\,\lambda_{\max}(A + A^*)$; this quantity is also called the *logarithmic norm* of A. We will now review some areas where it may be of interest to approximate the quantities $\nu(A^{-1})$, $\mu(A^{-1})$, or $\omega(A^{-1})$.

The first area we consider is the study of the convergence of iterative methods for linear systems. Consider the linear system $Ax = b$ with x_0 as initial guess, $r_0 = b - Ax_0$ as corresponding initial residual, and the Krylov space $\mathcal{U}_k = \mathcal{K}(A, r_0)$ as search space. In minimum residual methods, such as MINRES for symmetric A or GMRES for unsymmetric A, one approximates x by

$$x \approx x_k = (AU_k)^+ b := (U_k^* A^* A U_k)^{-1} U_k^* A^* b,$$

where we assume that $A\mathcal{U}_k$ is of full rank k. This choice minimizes the residual over the search space: $x_k = \mathrm{argmin}_{y_k \in \mathcal{U}_k} \|b - Ay_k\|$. Eiermann and Ernst [4, Section 6.1.3], [3, Cor. 6.2] show that for this minimum residual

$$\|r_k^{\mathrm{MR}}\| := \|b - Ax_k\|$$

we have the upper bound

$$\frac{\|r_k^{\mathrm{MR}}\|}{\|r_0\|} \leq \left(1 - \nu(A)\,\nu(A^{-1})\right)^{k/2}.$$

Note in particular that the right-hand side is independent of the choice of the initial guess and the resulting Krylov search space.

Moreover, we have the following property. Since $x_k \in \mathcal{K}_k(A,b) = \operatorname{span}\{b, Ab, \ldots, A^{k-1}b\}$, we can write $x_k = q_{k-1}(A)\,b$ for a certain polynomial q_{k-1} of degree $k-1$, and $r_k = p_k(A)\,b$ for a certain polynomial p_k of degree k. Then the zeros ζ of the *GMRES polynomial* p_k satisfy $\zeta^{-1} \in \mathrm{FOV}(A^{-1})$, see [8] and also below.

In fact, these zeros are the *harmonic Ritz values* [10, 5]. The harmonic Rayleigh–Ritz approach, which was introduced in [10], is is a popular tool for the numerical computation of interior eigenvalues close to a given target τ. Note that these eigenvalues are exterior eigenvalues of the shifted and inverted matrix $(A-\tau I)^{-1}$, where I denotes the identity matrix. For subspace methods it is often comparatively easy to compute exterior eigenvalues. Given a search space $\mathcal{U}_k$ for the eigenvector, this suggests to determine (or extract) approximate eigenpairs (θ, u), where $u \in \mathcal{U}_k$, with the Galerkin condition

$$(A-\tau I)^{-1}u - (\theta-\tau)^{-1}u \perp \widetilde{\mathcal{U}}_k.$$

Here $\widetilde{\mathcal{U}}_k$ is a test space; to avoid working with the inverse of a large sparse matrix the harmonic Rayleigh–Ritz approach takes $\widetilde{\mathcal{U}}_k = (A-\tau I)^*(A-\tau I)\,\mathcal{U}_k$. In this case the subspace extraction is determined by the projected (and hence low-dimensional) generalized eigenvalue problem

$$U_k^*(A-\tau I)^*(A-\tau I)\,U_k c = (\theta-\tau)\,U_k^*(A-\tau I)^* U_k c,$$

where we write $u = U_k c$, with $c \in \mathbb{C}^k$. The pair $(\theta, u) = (\theta, U_k c)$ is called a harmonic Ritz pair. For eigenpair approximations, one is interested in the pair with the harmonic Ritz value θ closest to τ, see, e.g., [13]. Given a harmonic Ritz vector u, the corresponding harmonic Ritz value is given by

$$\frac{u^*(A-\tau I)^*(A-\tau I)\,u}{u^*(A-\tau I)^* u}.$$

The roots of the GMRES polynomial are the harmonic Ritz values with target $\tau = 0$ [10, 5]. By the equalities

$$\mathrm{FOV}(A^{-1}) = \left\{ \frac{x^*A^{-1}x}{x^*x} : x \neq 0 \right\} = \left\{ \frac{y^*A^*y}{y^*A^*Ay} : y \neq 0 \right\}$$

we see that the inverses of these harmonic Ritz values are in $\mathrm{FOV}(A^{-1})$. These equalities may also be interpreted as: $\mathrm{FOV}(A^{-1})$ is the set of the inverses of all harmonic Ritz values that may be obtained after one step of the Arnoldi iteration (see [15, p. 166] for a similar statement about $\mathrm{FOV}(A)$ and Ritz values). We will come back to the relation of harmonic Ritz values and the field of values in Section 2.3.

For another area of application, we consider the linear differential-algebraic equation with constant coefficients

$$Au'(t) = u(t), \qquad u(0) = u_0.$$

If A is expensive to invert, this equation may be considered "effectively implicit" in the sense that it is computationally not easy to explicitly express $u'(t)$ in terms of $u(t)$. The numerical abscissa of A^{-1} determines the behavior of $\|e^{tA^{-1}}\|$ as $t \to 0$ (see, e.g., [15, (14.2)]):

$$\frac{d}{dt}\,\|e^{tA^{-1}}\|_{t=0} = \omega(A^{-1}).$$

In fact, we also have $\|e^{tA^{-1}}\| \leq e^{t\omega(A^{-1})}$ for all $t \geq 0$ [15, Thm. 17.1].

We can also use $\mathrm{FOV}(A^{-1})$ for estimates for the norm of the matrix inverse and the condition number. We have

$$\mu(A^{-1}) \leq \|A^{-1}\| \leq 2\,\mu(A^{-1})$$

(and a similar result for powers of A^{-1}, cf. [15, (17.6)]). Combined with an estimate for $\|A\|$, for instance using $\mu(A) \leq \|A\| \leq 2\,\mu(A)$, this gives an approximation for the condition number; see also the following subsection.

Finally, the field of values of A and of A^{-1} may be used to determine inclusion regions for matrix eigenvalues. Particularly, Manteuffel and Starke [8] mention the inclusion region

$$\Lambda(A) \subseteq \mathrm{FOV}(A) \;\cap\; \frac{1}{\mathrm{FOV}(A^{-1})}.$$

We will comment on this inclusion in the examples in Section 2.3.

2.2. Bounds and relations. Before considering the practical approximation of the field of values of A^{-1} in the next subsection, we look at some bounds for $\nu(A^{-1})$ and $\mu(A^{-1})$, and their relations with the quantities $\mu(A)$, $\nu(A)$, $\|A\|$, and $\|A^{-1}\|$. Eiermann and Ernst [3] note that

$$\nu(A^{-1}) = \min_{x\neq 0} \left|\frac{x^*A^{-1}x}{x^*x}\right| = \min_{w\neq 0} \left|\frac{w^*A^*w}{w^*w}\right| \cdot \left|\frac{w^*w}{w^*A^*Aw}\right| \geq \frac{\nu(A)}{\|A\|^2}. \tag{2.1}$$

Using the same reasoning we can also derive

$$\nu(A^{-1}) \leq \min\left\{ \frac{\nu(A)}{\sigma_{\min}^2(A)}\,,\; \frac{\mu(A)}{\|A\|^2} \right\}.$$

In particular, we have from this

$$\|A\| \leq (\nu(A^{-1}))^{-1}\, \frac{\mu(A)}{\|A\|} \leq (\nu(A^{-1}))^{-1}$$

and therefore $\|A\| \leq \min\{2\,\mu(A),\; \nu(A^{-1})^{-1}\}$.

An expression similar to (2.1) for the numerical radius is

$$\mu(A^{-1}) = \max_{x\neq 0} \left|\frac{x^*A^{-1}x}{x^*x}\right| = \max_{w\neq 0} \left|\frac{w^*A^*w}{w^*w}\right| \cdot \left|\frac{w^*w}{w^*A^*Aw}\right| \leq \frac{\mu(A)}{\sigma_{\min}^2(A)}.$$

From the same expressions we also get

$$\mu(A^{-1}) \geq \max\left\{ \frac{\nu(A)}{\sigma_{\min}^2(A)}\,,\; \frac{\mu(A)}{\|A\|^2} \right\}.$$

From a different perspective the bounds in this subsection, together with bounds such as $\mu(A^{-1}) \leq \|A^{-1}\| \leq 2\,\mu(A^{-1})$, imply bounds for $\|A\|$, $\|A^{-1}\|$ and the condition number $\kappa(A) = \|A\|\,\|A^{-1}\|$ in terms of $\mathrm{FOV}(A)$ and $\mathrm{FOV}(A^{-1})$.

2.3. Two approximation methods. To approximate the field of values of A^{-1}, note that we have

$$\mathrm{FOV}(A^{-1}) \supseteq \mathrm{FOV}(\widetilde{U}_k^* A^{-1} \widetilde{U}_k)$$

for any $n \times k$ matrix $\widetilde{U}_k$ with orthonormal columns that form a basis for a space $\widetilde{\mathcal{U}}_k$. To avoid matrix inversion of a large matrix, one idea is to take $\widetilde{\mathcal{U}}_k = A\mathcal{U}_k$. We can compute the resulting approximation efficiently as follows. If $\underline{H}_k = Q_k R_k$ is the reduced QR-decomposition of $\underline{H}_k$, then $AU_kR_k^{-1} = U_{k+1}\underline{H}_kR_k^{-1} = U_{k+1}Q_k$ has orthonormal columns, and

$$\mathrm{FOV}(A^{-1}) \supseteq \mathrm{FOV}(R_k^{-*}U_k^*A^*A^{-1}AU_kR_k^{-1}) = \mathrm{FOV}(R_k^{-*}H_k^*R_k^{-1}). \tag{2.2}$$

We note that Manteuffel and Starke [8] also suggest this approximation via a slightly different derivation. An advantage of expression (2.2) compared with [8] is that from (2.2) it

is clear that this approximation to $\mathrm{FOV}(A^{-1})$ is itself a field of values of a low-dimensional matrix $R_k^{-*}H_k^*R_k^{-1}$. (In [8], this approximation is described as the intersection of strips in the complex plane determined by the minimal and maximal eigenvalues of *generalized* eigenvalue problems.)

In the following we will assume that H_k is invertible. The eigenvalues of

$$\widetilde{H}_k := (U_k^*A^*U_k)^{-1}U_k^*A^*AU_k = H_k^{-*}R_k^*R_k$$

are the *harmonic Ritz values* of A with respect to search space $\mathcal{U}_k$ and target $\tau = 0$ (see [10] and Section 2.1). It is not difficult to prove that the eigenvalues of this matrix are the inverses of the eigenvalues of $R_k^{-*}H_k^*R_k^{-1}$. Since the eigenvalues of $\widetilde{H}_k$ are the harmonic Ritz values, our conclusion is that after k steps we know that $\mathrm{FOV}(A^{-1})$ contains the convex hull of the inverses of the harmonic Ritz values.

We now present a second, alternative, approach to approximate $\mathrm{FOV}(A^{-1})$:

$$\mathrm{FOV}(A^{-1}) \supseteq \mathrm{FOV}(U_k^*A^{-1}U_k) \approx \mathrm{FOV}(H_k^{-1}). \tag{2.3}$$

To derive this approximation we rewrite (1.2):

$$A^{-1}U_k = U_kH_k^{-1} - h_{k+1,k}A^{-1}u_{k+1}e_k^*H_k^{-1},$$

so that

$$U_k^*A^{-1}U_k = H_k^{-1} - h_{k+1,k}U_k^*A^{-1}u_{k+1}e_k^*H_k^{-1}. \tag{2.4}$$

Discarding the last term on the right-hand side gives the approximation $U_k^*A^{-1}U_k \approx H_k^{-1}$. In fact, the field of values of H_k^{-1} is the field of values of a projected nearby matrix:

$$\mathrm{FOV}(H_k^{-1}) = \mathrm{FOV}(U_k^*(A^{-1} + E_k)\,U_k)$$

where the *backward error*

$$E_k = h_{k+1,k}(A^{-1}u_{k+1})(e_k^*H_k^{-1}U_k^*) \tag{2.5}$$

is a (hopefully small) rank-one update matrix. In particular $\|E_k\|$, will be small if $\mathcal{U}_k$ is almost an invariant subspace, which implies that $|h_{k+1,k}|$ is small. Or, in terms of a multiplicative (relative) perturbation,

$$\mathrm{FOV}(H_k^{-1}) = \mathrm{FOV}(U_k^*A^{-1}(I_k + F_k)\,U_k),$$

where

$$F_k = h_{k+1,k}u_{k+1}(e_k^*H_k^{-1}U_k^*).$$

We can be a bit more precise. From (2.4) we have

$$y^*H_k^{-1}y = y^*(U_k^*A^{-1}U_k)\,y + h_{k+1,k}(y^*U_k^*A^{-1}u_{k+1})(e_k^*H_k^{-1}y)$$

from which we see that for all $z \in \mathrm{FOV}(H_k^{-1})$ we have

$$\mathrm{dist}(z,\ \mathrm{FOV}(U_k^*A^{-1}U_k)) \leq |h_{k+1,k}|\,\|U_k^*A^{-1}u_{k+1}\|\,\|H_k^{-*}e_k\|.$$

In other words: every point of $\mathrm{FOV}(H_k^{-1})$ is at most $|h_{k+1,k}|\,\|U_k^*A^{-1}u_{k+1}\|\,\|H_k^{-*}e_k\|$ away from $\mathrm{FOV}(U_k^*A^{-1}U_k)$, a projected field of values of A^{-1} which is included in the sought set $\mathrm{FOV}(A^{-1})$.

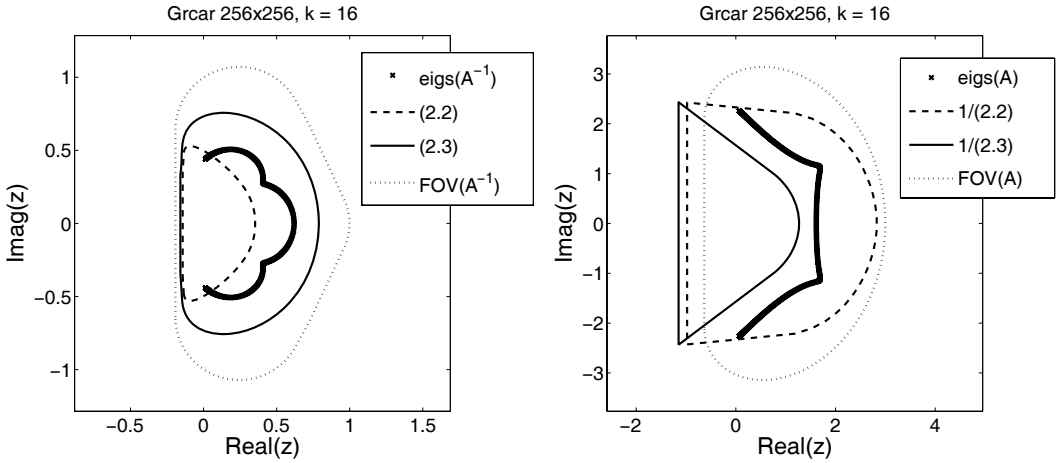
We point out that there is another way to derive the second approximation (2.3). Suppose that in the first method (see (2.2)) one takes the QR-decomposition of H_k instead of that of $\underline{H}_k$. Then

$$\begin{aligned}\mathrm{FOV}(A^{-1}) &\approx \mathrm{FOV}(R_k^{-*}H_k^*R_k^{-1}) = \mathrm{FOV}(Q_k^*R_k^{-1}) \\ &= \mathrm{FOV}(Q_kQ_k^*R_k^{-1}Q_k^*) = \mathrm{FOV}(R_k^{-1}Q_k^{-1}) = \mathrm{FOV}(H_k^{-1}).\end{aligned}$$

In summary, when we compare (2.2) and (2.3), the advantage of the former is that this approximation is a proper inclusion. However, a strength of (2.3) over (2.2) is that this method employs a projection that works solely with $\mathcal{U}_k$: (2.3) approximation a projection of A^{-1} onto the space $\mathcal{U}_k$, and not onto the space $A\mathcal{U}_k$. The space $A\mathcal{U}_k$ can be seen as biased, as by the multiplication with A the smallest eigenmodes will have been (greatly) reduced. Therefore, we expect that (2.3) may in practice yield the best approximation, especially if $|h_{k+1,k}|\,\|U_k^*A^{-1}u_{k+1}\|\,\|H_k^{-*}e_k\|$ is reasonably small.

We now continue with a few numerical examples. The first example shows another possible advantage of the approximation $\mathrm{FOV}(H_k^{-1})$ over $\mathrm{FOV}(R_k^{-*}H_k^*R_k^{-1})$.

EXAMPLE 2.1. In Figure 2.1(a) the eigenvalues and FOV (dotted) of A^{-1} are plotted, where A is the 256×256 `grcar` matrix and $\mathcal{U}$ is a 16-dimensional Krylov space generated using a random starting vector. We also plot two approximations to $\mathrm{FOV}(A^{-1})$: $\mathrm{FOV}(R_k^{-*}H_k^*R_k^{-1})$ ((2.2), dash), and $\mathrm{FOV}(H_k^{-1})$ ((2.3), solid). Although it is not guaranteed that $\mathrm{FOV}(H_k^{-1})$ is a subset of $\mathrm{FOV}(A^{-1})$, it is the case here; moreover, $\mathrm{FOV}(H_k^{-1})$ is a much better approximation to $\mathrm{FOV}(A^{-1})$ than is $\mathrm{FOV}(R_k^{-*}H_k^*R_k^{-1})$.



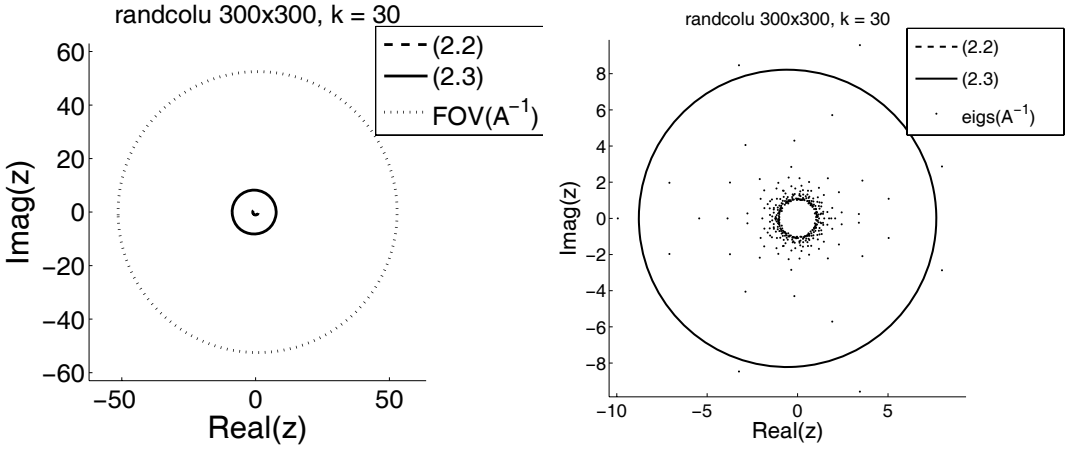
(a) (b)

FIG. 2.1. *(a) Data for A^{-1}: eigenvalues, field of values, and the two approximations of the FOV given by* (2.2) *($A^{-1}|_{A\mathcal{U}}$, dash) and* (2.3) *(FOV(H_k^{-1}), solid, which is an approximation to the FOV of $A^{-1}|_{\mathcal{U}}$). (b) Eigenvalues and FOV of A, and, in addition, $1/F$, where F is the approximation to the* $\mathrm{FOV}(A^{-1})$ *according to* (2.2) *and* (2.3)*, respectively.*

In Figure 2.1(b) we take the inverses of the sets determined by (2.2) and (2.3) and also display the eigenvalues and FOV of A. Recall the eigenvalue inclusion region $\Lambda(A) \subseteq \mathrm{FOV}(A) \cap \frac{1}{\mathrm{FOV}(A^{-1})}$. In Figure 2.1(b), the approximations (2.2) and (2.3) suggest that $\Lambda(A)$ should be on the *outside* of the dashes, respectively solid line. We see that this is correct for the approximation based on (2.3) but incorrect for the one based on (2.2). Therefore, the approximation $\mathrm{FOV}(H_k^{-1})$ to $\mathrm{FOV}(A^{-1})$ may also yield better spectral inclusion regions. ⊘

EXAMPLE 2.2. The next test case, which we denote by `randcolu`, is generated by the MATLAB command `A = gallery('randcolu', 300)`. This is a random 300×300 matrix

with columns of unit two-norm, with random singular values whose squares are from a uniform distribution. The initial vector is random from which we generate a Krylov search space $\mathcal{U}$ of dimension 30.

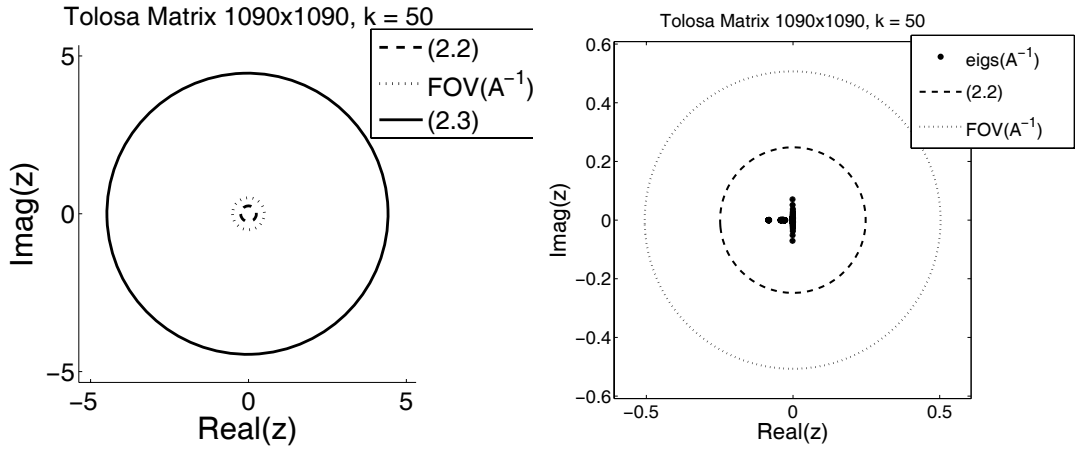


(a) (b)

FIG. 2.2. *(a)* FOV(A^{-1}) *where* A *is the* 300×300 `randcolu` *matrix and approximations* (2.2) *and* (2.3) *based on a 30-dimensional Krylov space. (b): Zoom-in of (a), also showing the eigenvalues of* A^{-1}.

In Figure 2.2, we see FOV(A^{-1}) as well as the approximations (2.2) and (2.3). The (b) figure zooms in on part of (a); we see that (2.3) is considerably better than (2.2), in particular (2.3) includes most eigenvalues of A^{-1} in contrast to (2.2); however, both approximations are a significant underestimation of FOV(A^{-1}). ⊘

EXAMPLE 2.3. In the final example we consider the `tolosa` matrix [9] of dimension 1090. We take a 50-dimensional Krylov space starting with a random initial vector.



(a) (b)

FIG. 2.3. *(a)* FOV(A^{-1}) *where* A *is the* 1090×1090 `tolosa` *matrix and approximations* (2.2) *and* (2.3) *based on a 50-dimensional Krylov space. (b): Zoom-in on (a), also showing the eigenvalues of* A^{-1}.

In Figure 2.3(a) we show FOV(A^{-1}) and the approximations (2.2) and (2.3). Figure 2.3(b) zooms in on part of the (a) figure and also show the eigenvalues of A^{-1}. We see that in this case (2.2) seems a better approximation than (2.3). In fact, (2.3) is a crude overestimation while (2.2) is an underestimation. ⊘

2.4. A corrected approximation. Inspired by the relation (2.4) between $U_k^* A^{-1} U_k$ and H_k^{-1}, we can also approximation FOV(A^{-1}) by computing the FOV of the "corrected matrix" $H_k^{-1} - h_{k+1,k} U_k^* \widetilde{y}_{k+1} e_k^* H_k^{-1}$, where $\widetilde{y}_{k+1}$ is an approximation to $y_{k+1} := A^{-1} u_{k+1}$.

We may approximate $A^{-1}u_{k+1}$ by a second Arnoldi process. For instance, this may be a sensible idea when the search space $\mathcal{U}_k$ reaches a dimension so that further expansion of the space would become computationally unattractive because of basis vectors storage or orthogonalization costs, or when H_k^{-1} reaches a size so that the computation of its FOV can no longer be considered cheap. We will discuss a few more details on a corrected approximation in Section 3.4 in the context of pseudospectra computation.

3. Approximating the pseudospectrum of a matrix inverse. We now switch our attention to the pseudospectrum. This valuable tool for understanding the behavior of nonnormal matrices has also been used to study convergence of iterative methods. We recall the definition of the ε-pseudospectrum [15] for $\varepsilon \geq 0$:

$$\begin{aligned}\Lambda_\varepsilon(A) &= \{\, z : \sigma_{\min}(A - zI) \leq \varepsilon \,\} \\ &= \{\, z : \|(A - zI)^{-1}\| \geq \varepsilon^{-1} \,\} \\ &= \{\, z : (A - zI + E) \text{ singular for an } E \text{ with } \|E\| \leq \varepsilon \,\} \\ &= \{\, z : (A + E)\, x = zx \text{ for } \|E\| \leq \varepsilon \text{ and } \|x\| = 1 \,\}.\end{aligned}$$

(Note that in contrast to [15], but in line with earlier publications, we prefer the "closed set" definition.) Computing the smallest singular value for a large number of (grid points) $z \in \mathbb{C}$ values is a very expensive task. Therefore, motivated by

$$U_{k+1}^*(A - zI)\, U_k = \underline{H}_k - z\underline{I},$$

for all $z \in \mathbb{C}$, where $\underline{I}$ denotes the identity with an additional row of zeros, Toh and Trefethen [14] and Wright and Trefethen [17] approximate $\sigma_{\min}(A - zI)$ by $\sigma_{\min}(\underline{H}_k - z\underline{I})$, where

$$\sigma_{\min}(A - zI) \leq \sigma_{\min}(\underline{H}_k - z\underline{I}).$$

This means that for all $\varepsilon \geq 0$

$$\Lambda_\varepsilon(\underline{H}_k) \subseteq \Lambda_\varepsilon(A).$$

We are now interested in approximating the pseudospectrum of A^{-1}. First, we look at some applications where such information may be useful.

3.1. Applications. Pseudospectra are a popular tool for understanding the behavior of nonnormal matrices, as well as estimating norms of matrix powers, matrix polynomials, and more general matrix functions. Recall the Dunford–Taylor integral representation for a matrix function (see, e.g., [15, (14.9)])

$$f(A) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - A)^{-1} dz$$

where Γ is the boundary curve of a piecewise smooth, bounded region containing the spectrum of A, assuming that f is analytic in the region and continuous on its closure.

Let $m \in \mathbb{N}$. From the integral representation

$$A^m = \frac{1}{2\pi i} \int_\Gamma z^m\, (zI - A)^{-1} dz$$

we get, if we take the boundary curve of the ε-pseudospectra $\Lambda_\varepsilon(A)$ for Γ,

$$\|A^m\| \leq \frac{\ell(\Lambda_\varepsilon(A))}{2\pi\varepsilon}\, \rho_\varepsilon^m(A)$$

for every $\varepsilon > 0$ (cf., e.g., [15, (14.10)]), where the ε-pseudospectral radius is given by

$$\rho_\varepsilon(A) = \max_{z \in \Lambda_\varepsilon(A)} |z|$$

and $\ell(\Gamma)$ denotes the arc length of the curve Γ.

We now give another bound using the pseudospectra of A^{-1}. If the region bounded by Γ does not contain the origin, we have

$$A^m = (A^{-1})^{-m} = \frac{1}{2\pi i} \int_\Gamma z^{-m} \, (zI - A^{-1})^{-1} dz.$$

Taking the boundary of $\Lambda_\varepsilon(A^{-1})$ for Γ, we get for every $\varepsilon > 0$ such that $0 \notin \Lambda_\varepsilon(A^{-1})$

$$\|A^m\| \leq \frac{\ell(\Lambda_\varepsilon(A^{-1}))}{2\pi\varepsilon} \max_{z \in \Lambda_\varepsilon(A^{-1})} |z|^{-m}.$$

An extension for matrix polynomials p is given by (cf., e.g., [15, (14.10)])

$$\|p(A)\| \leq \frac{\ell(\Lambda_\varepsilon(A))}{2\pi\varepsilon} \max_{z \in \Lambda_\varepsilon(A)} |p(z)|.$$

We can also derive a bound in terms of $\Lambda_\varepsilon(A^{-1})$ as follows. Let m be the degree of p and let $\widetilde{p}$ be the polynomial such that $p(z) = z^m \widetilde{p}(z^{-1})$. Then from

$$p(A) = A^m \, \frac{1}{2\pi i} \int_\Gamma \widetilde{p}(z) \, (zI - A^{-1})^{-1} dz$$

we get the bound

$$\|p(A)\| \leq \frac{\ell(\Lambda_\varepsilon(A^{-1}))}{2\pi\varepsilon} \, \|A^m\| \max_{z \in \Lambda_\varepsilon(A^{-1})} |\widetilde{p}(z)|.$$

If the region bounded by Γ does not contain the origin then

$$p(A) = A^m \, \frac{1}{2\pi i} \int_\Gamma \widetilde{p}(z^{-1}) \, (zI - A)^{-1} dz$$

from which it follows that if $0 \notin \Lambda_\varepsilon(A)$ then we have

$$\|p(A)\| \leq \frac{\ell(\Lambda_\varepsilon(A))}{2\pi\varepsilon} \, \|A^m\| \max_{z \in \Lambda_\varepsilon(A)} |\widetilde{p}(z^{-1})|.$$

Here, $\|A^m\|$ may be bounded from above by $2\,(\mu(A))^m$ [15, (17.6)].

Finally, we mention that $\Lambda_\varepsilon(A^{-1})$ gives information about the behavior of $\|e^{tA^{-1}}\|$ (see [15, Ch. 15]): for all $t \geq 0$ and $\varepsilon > 0$ we have

$$\frac{\alpha_\varepsilon(A^{-1})}{\varepsilon} \leq \|e^{tA^{-1}}\| \leq \frac{\ell(\Lambda_\varepsilon(A^{-1}))\, e^{t\alpha_\varepsilon(A^{-1})}}{2\pi\varepsilon}.$$

3.2. A relation. Naturally, we have the relation $\Lambda(A^{-1}) = (\Lambda(A))^{-1}$ between the spectra of A and A^{-1}. We now give a extension of this statement for the pseudospectra of A and of A^{-1}.

PROPOSITION 3.1. *Let A be invertible and $z \neq 0$. Then we have the following implications:*

$$z \in (\Lambda_{(\sigma_{\min}(A)/|z|)\,\varepsilon}(A))^{-1} \;\Rightarrow\; z \in \Lambda_\varepsilon(A^{-1}) \;\Rightarrow\; z \in (\Lambda_{(\|A\|/|z|)\,\varepsilon}(A))^{-1}.$$

Proof. We will use $zI - A^{-1} = zA^{-1}(A - z^{-1}I)$. Suppose $z \in \Lambda_\varepsilon(A^{-1})$. Since

$$\sigma_{\min}(zI - A^{-1}) \geq |z| \, \sigma_{\min}(A^{-1}) \, \sigma_{\min}(A - z^{-1}I) = (|z|/\|A\|) \cdot \sigma_{\min}(A - z^{-1}I)$$

we have $\sigma_{\min}(A - z^{-1}I) \leq (\|A\|/|z|) \, \sigma_{\min}(zI - A^{-1}) \leq (\|A\|/|z|)\, \varepsilon$ and therefore $z^{-1} \in \Lambda_{(\|A\|/|z|)\,\varepsilon}(A)$. For the first implication we use the inequality

$$\sigma_{\min}(zI - A^{-1}) \leq |z| \, \sigma_{\max}(A^{-1}) \, \sigma_{\min}(A - z^{-1}I) = (|z|/\sigma_{\min}(A)) \cdot \sigma_{\min}(A - z^{-1}I)$$

from which follows that if $z^{-1} \in \Lambda_{(\sigma_{\min}(A)/|z|)\,\varepsilon}(A)$ then $z \in \Lambda_\varepsilon(A^{-1})$. ∎

3.3. Three approximation methods. Let us consider approximation methods for $\Lambda_\varepsilon(A^{-1})$. First, note that the condition number of a simple eigenvalue λ^{-1} of A^{-1} is the same as that of the corresponding eigenvalue λ of A:

$$\kappa(\lambda^{-1}) = |y^*x|^{-1} = \kappa(\lambda),$$

where κ stands for the condition number, and x and y are the right and left eigenvectors, respectively, corresponding to eigenvalue λ^{-1} of A^{-1} (and to eigenvalue λ of A).

Hence, if all eigenvalues are simple, we can expect that for $\varepsilon \downarrow 0$, the pseudospectrum $\Lambda_\varepsilon(A^{-1})$ consists of disks of radius $\varepsilon\,\kappa(\lambda)$ around the eigenvalues λ^{-1}, just as $\Lambda_\varepsilon(A)$ asymptotically consists of disks of radius $\varepsilon\,\kappa(\lambda)$ around the eigenvalues λ. However, this is only an asymptotic statement; the pseudospectrum for any chosen $\varepsilon > 0$ may generally be much less simple than the union of these disks, in particular if A has multiple eigenvalues or if A is (highly) nonnormal. Therefore, it may be of interest to approximate $\Lambda_\varepsilon(A^{-1})$ without computing A^{-1}, or its eigendata, first. This amounts to approximating $\sigma_{\min}(A^{-1} - zI)$, for a large number of grid points z.

We present three approximation methods by means of projection onto subspaces; each method makes a particular choice for the right and left $n \times k$ projection matrices V_k and W_k with orthonormal columns in the inequality

$$\sigma_{\min}(A^{-1} - zI) \le \sigma_{\min}(W_k^*(A^{-1} - zI)V_k).$$

Hereby, we bear in mind that we would like to avoid an action with A^{-1}.

As in the previous section, let $\underline{H}_k = Q_k R_k$ be the reduced QR-decomposition of $\underline{H}_k$. First, we may choose $V_k = AU_kR_k^{-1}$ and $W_k = U_k$ to get

$$\begin{aligned}\sigma_{\min}(A^{-1} - zI) &\le \sigma_{\min}(U_k^*(A^{-1} - zI)AU_kR_k^{-1}) \\ &= \sigma_{\min}((I_k - zH_k)R_k^{-1}).\end{aligned} \tag{3.1}$$

Second, we may take $V_k = W_k = AU_kR_k^{-1}$ to get

$$\begin{aligned}\sigma_{\min}(A^{-1} - zI) &\le \sigma_{\min}(R_k^{-*}U_k^*A^*(A^{-1} - zI)AU_kR_k^{-1}) \\ &= \sigma_{\min}(R_k^{-*}H_k^*R_k^{-1} - zI).\end{aligned} \tag{3.2}$$

Finally, we may choose $V_k = W_k = U_k$. Since $U_k^*A^{-1}U_k$ is not practical we approximate

$$\begin{aligned}\sigma_{\min}(A^{-1} - zI) &\le \sigma_{\min}(U_k^*A^{-1}U_k - zI) \\ &\approx \sigma_{\min}(H_k^{-1} - zI).\end{aligned} \tag{3.3}$$

Let us now analyze the quality of (3.3) in some more detail. First, making use of backward error (2.5) we can show as in the previous section that $H_k^{-1} - zI$ can be considered as a projection of a perturbation of $A^{-1} - zI$. Moreover, using (1.2) we get

$$I_k = (U_k^*A^{-1}U_{k+1})\,\underline{H}_k.$$

From this equality, and using

$$U_k^*U_{k+1}\underline{H}_k = H_k,$$

we get

$$I_k - zH_k = (U_k^*(A^{-1} - zI)\,U_{k+1})\,\underline{H}_k,$$

or

$$H_k^{-1} - zI_k = (U_k^*(A^{-1} - zI)\,U_{k+1})\begin{bmatrix} I \\ h_{k+1,k}e_k^*H_k^{-1} \end{bmatrix}.$$

We can compare the minimal singular values of $H_k^{-1} - zI_k$ and $U_k^*(A^{-1} - zI)\,U_{k+1}$ as follows. On the one hand, we have

$$\begin{aligned}\sigma_{\min}(H_k^{-1} - zI_k) &\le \sigma_{\min}(U_k^*(A^{-1} - zI)\,U_{k+1})\,\sigma_{\max}\left(\begin{bmatrix} I \\ h_{k+1,k}e_k^*H_k^{-1}\end{bmatrix}\right)\\ &\le \sigma_{\min}(U_k^*(A^{-1} - zI)\,U_{k+1})\,\sqrt{1 + |h_{k+1,k}|^2\,\|H_k^{-*}e_k\|^2}.\end{aligned}$$

On the other, we have

$$\begin{aligned}\sigma_{\min}(H_k^{-1} - zI_k) &\ge \sigma_{\min}(U_k^*(A^{-1} - zI)\,U_{k+1})\,\sigma_{\min}\left(\begin{bmatrix} I \\ h_{k+1,k}e_k^*H_k^{-1}\end{bmatrix}\right)\\ &\ge \sigma_{\min}(U_k^*(A^{-1} - zI)\,U_{k+1}).\end{aligned}$$

So if $|h_{k+1,k}|\,\|H_k^{-*}e_k\|$ is small, then we conclude that $\sigma_{\min}(H^{-1} - zI)$ is a good approximation to $\sigma_{\min}(U_k^*(A^{-1} - zI)\,U_{k+1})$. Similar to the comparison of (2.2) and (2.3) in the previous section, an advantage of the approach (3.3) over (3.1) and (3.2) is that it approximates a projection of $A^{-1} - zI$ onto the space $\mathcal{U}_k$ instead of onto the space $A\mathcal{U}_k$. Therefore, (3.3) may practically give the best approximation, as long as $|h_{k+1,k}|\,\|H_k^{-*}e_k\|$ is not too large.

We note that the methods presented here are in some sense the opposite compared to [14, Section 5], where $\Lambda_\varepsilon(A)$ is approximated by an Arnoldi process on A^{-1}; in this paper we assume that this is computationally unattractive. See also Simoncini and Gallopoulos [11] for approaches based on rational Krylov.

In addition to the three methods presented in this subsection, we will consider a corrected scheme in the next subsection, which employs a second Krylov space.

3.4. A corrected approximation. We now present another approximation method for $\Lambda_\varepsilon(A^{-1})$ inspired by work of Bekas, Gallopoulos, and Simoncini [11, 2], who worked directly with a projection of $(A - zI)^{-1}$ to get a better approximation to the resolvent norm, and hence the pseudospectra. For our situation, note that

$$(A^{-1} - zI)^{-1} = (I - zA)^{-1}A.$$

We start with

$$(I - zA)^{-1}AU_k = (I - zA)^{-1}U_kH_k + h_{k+1,k}(I - zA)^{-1}u_{k+1}e_k^*. \tag{3.4}$$

Since

$$(I - zA)\,U_k = U_k(I - zH_k) - h_{k+1,k}zu_{k+1}e_k^*$$

we have

$$(I - zA)^{-1}U_k = U_k(I - zH_k)^{-1} + h_{k+1,k}z\,(I - zA)^{-1}u_{k+1}e_k^*(I - zH_k)^{-1}.$$

Substituting this into (3.4) gives

$$(I - zA)^{-1}AU_k = U_k(H_k^{-1} - zI)^{-1} + h_{k+1,k}(I - zA)^{-1}u_{k+1}e_k^*(z(H_k^{-1} - zI)^{-1} + I).$$

We have

$$z\,(H_k^{-1} - zI)^{-1} + I = (H_k^{-1} - zI)^{-1}(zI + H_k^{-1} - zI) = (H_k^{-1} - zI)^{-1}H_k^{-1} = (I - zH_k)^{-1}.$$

Our conclusion is that

$$\begin{aligned}U_k^*(I - zA)^{-1}AU_k &= (H_k^{-1} - zI)^{-1} + h_{k+1,k}U_k^*(I - zA)^{-1}u_{k+1}e_k^*(I - zH_k)^{-1}\\ &= (H_k + h_{k+1,k}U_k^*(I - zA)^{-1}u_{k+1}e_k^*)(I - zH_k)^{-1}.\end{aligned}$$

Let $y_{k+1}(z)$ be the solution to $(I - zA)^{-1}u_{k+1}$. The key idea is now that $y_{k+1}(z)$ can be approximated by a second Krylov procedure, and that for $z \neq 0$ the Krylov spaces $\mathcal{K}_m(I - zA, u_{k+1})$ are identical and equal to $\mathcal{K}_m(A, u_{k+1})$, while for $z = 0$ we get $y_{k+1}(0) = u_{k+1}$.

This suggests another approximation to the pseudospectra of A^{-1} by considering the pseudospectra of

$$U_k^*(I - zA)^{-1}AU_k \approx (H_k + h_{k+1,k}U_k^*\widetilde{y}_{k+1}(z)e_k^*)(I - zH_k)^{-1}, \tag{3.5}$$

where $\widetilde{y}_{k+1}(z)$ is an approximation to $y_{k+1}(z)$, for instance one determined by the Full Orthogonalization Method (FOM):

$$\widetilde{y}_{k+1}(z) = V_m(I - zG_m)^{-1}e_1,$$

where $AV_m = V_mG_m + g_{m+1,m}v_{m+1}e_m^*$ is a second Arnoldi decomposition of dimension m to form a basis for $\mathcal{K}_m(A, u_{k+1})$. We call (3.5) a corrected approximation since it involves for every $z \in \mathbb{C}$ a different rank-one update, that is computationally feasible since we only have to generate one extra low dimensional Krylov space $\mathcal{K}_m(A, u_{k+1})$.

EXAMPLE 3.2. In Figure 3.1, we compare the different approaches with the true pseudospectra of A^{-1} for A the 100×100 `grcar` matrix.

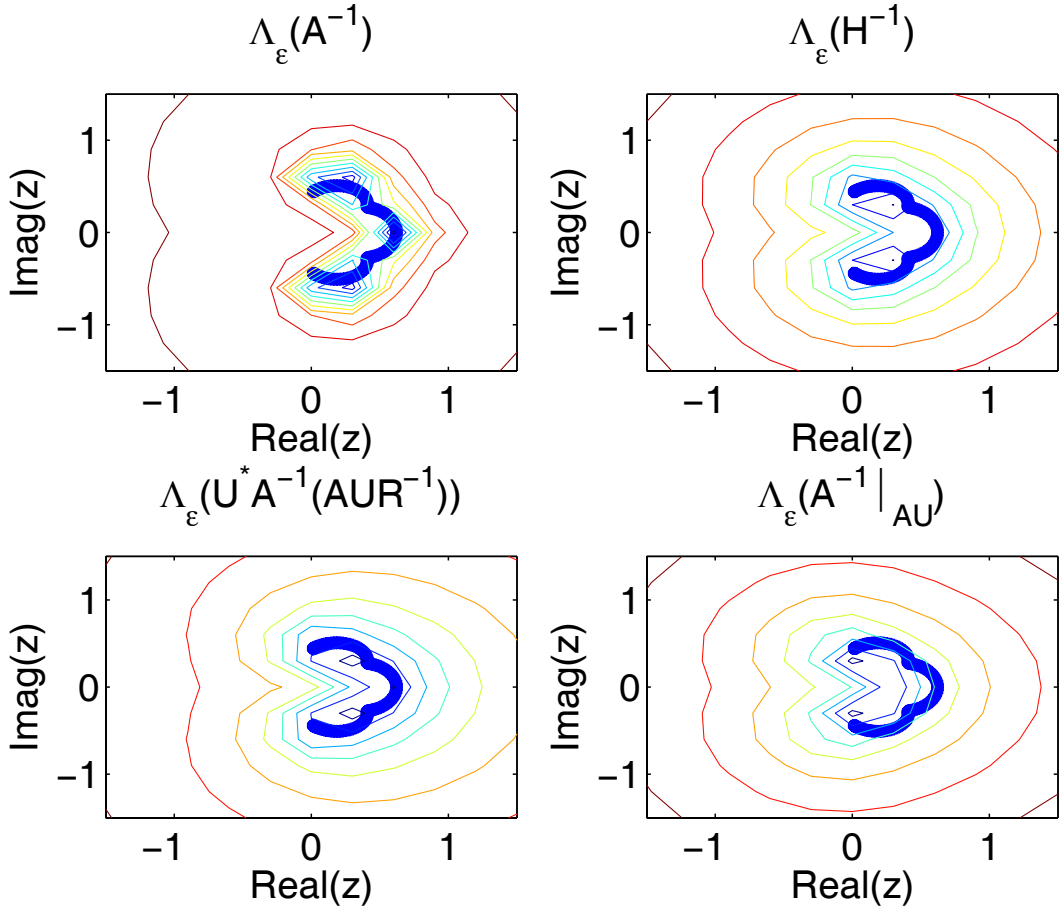


FIG. 3.1. $\Lambda_\varepsilon(A)$ *and various approximations for the* 100×100 `grcar` *matrix:* (3.3) *(top-right),* (3.1) *(bottom-left), and* (3.2) *(bottom-right). In all subfigures, the eigenvalues are also indicated.*

In addition to comparing the figures by sight, we take

$$\sum_{z \in \text{grid}} \left| {}^{10}\log(\sigma_{\min}(\text{approximation}(z)) - {}^{10}\log(\sigma_{\min}(A^{-1} - zI))) \right| \;/\; \#\text{gridpoints}$$

as an error measure. This error indicator is approximately $2.2 \cdot 10^{-4}$ for (3.1), $3.1 \cdot 10^{-4}$ for (3.2), and $2.1 \cdot 10^{-4}$ for (3.3). $\oslash$

4. Concluding remarks. We have approximated the field of values and pseudospectra of the inverse of a large sparse matrix by subspace methods, in particular the Arnoldi process. To this aim, we extended and gave alternatives for methods proposed by Manteuffel and Starke [8] for the field of values, and of Toh and Trefethen [14] for the pseudospectra.

The Arnoldi procedure can be used to give an approximation to the field of values and pseudospectra of A *and* of its inverse. As the Ritz values of A with respect to search space $\mathcal{U}_k$ are always in $\mathrm{FOV}(A)$, the inverses of the harmonic Ritz values are guaranteed to be in $\mathrm{FOV}(A^{-*})$. As we can use the *Ritz matrix* $H_k = U_k^* A U_k$ to approximate $\mathrm{FOV}(A)$, we can use the *harmonic Ritz matrix*, $\widetilde{H}_k = (U_k^* A^* U_k)^{-1} U_k^* A^* A U_k$, of which the harmonic Ritz values are the eigenvalues, to approximate $\mathrm{FOV}(A^{-1})$. However, H_k^{-1}, an approximation to the projected inverse matrix $U_k^* A^{-1} U_k$ may give more promising results.

We note that in principle we can use any search space $\mathcal{U}_k$ in the approximations. For instance, we may use the Jacobi–Davidson method [12] instead of the Arnoldi method. Of course, the analysis that makes use of (1.2) no longer holds but without further details we mention that one can derive similar results using $U_k^* A U_k$ instead of H_k and the residual matrix $R_k = AU_k - U_k H_k$ instead of $h_{k+1,k} u_{k+1} e_k^*$. However, Jacobi–Davidson generally focuses on a selected region in the complex plane of interest, which implies that the resulting approximate field of values may globally be of lower quality.

Related to this is the following comment. In [17] the Arnoldi procedure from [14] to approximate matrix pseudospectra was combined with an implicit restarted version of the Arnoldi process. This combination results in a high quality approximation of both a portion of the eigenvalues and the corresponding pseudospectra. On the other hand, the computations are generally (much) more time consuming and the resulting approximate pseudospectra may be locally excellent, but globally disappointing. The corrected schemes for FOV and pseudospectrum approximation as presented in this paper may be viewed as some sort of alternative to a restart.

REFERENCES

[1] W. E. Arnoldi, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.

[2] C. N. Bekas, E. Gallopoulos, and V. Simoncini, *Pseudospectra computation of large matrices*, Preprint, University of Minnesota, January 2004.

[3] M. Eiermann and O. G. Ernst, *Geometric aspects of the theory of Krylov subspace methods*, Acta Numer., 10 (2001), pp. 251–312.

[4] O. G. Ernst, *Minimal and orthogonal residual methods and their generalizations for solving linear operator equations*. Habilitationsschrift, TU Bergakademie Freiberg, Germany, 2000.

[5] S. Goossens and D. Roose, *Ritz and harmonic Ritz values and the convergence of FOM and GMRES*, Numer. Linear Algebra Appl., 6 (1999), pp. 281–293.

[6] C. R. Johnson, *Numerical determination of the field of values of a general complex matrix*, SIAM J. Numer. Anal., 15 (1978), pp. 595–602.

[7] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Stand., 45 (1950), pp. 255–282.

[8] T. A. Manteuffel and G. Starke, *On hybrid iterative methods for nonsymmetric systems of linear equations*, Numer. Math., 73 (1996), pp. 489–506.

[9] *The Matrix Market.* `http://math.nist.gov/MatrixMarket`, a repository for test matrices.

[10] C. C. Paige, B. N. Parlett, and H. A. van der Vorst, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Num. Lin. Alg. Appl., 2 (1995), pp. 115–133.

[11] V. Simoncini and E. Gallopoulos, *Transfer functions and resolvent norm approximation of large matrices*, Electron. Trans. Numer. Anal., 7 (1998), pp. 190–201.

[12] G. L. G. Sleijpen and H. A. van der Vorst, *A Jacobi–Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425.

[13] G. W. Stewart, *Matrix algorithms. Vol. II*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[14] K.-C. Toh and L. N. Trefethen, *Calculation of pseudospectra by the Arnoldi iteration*, SIAM J. Sci. Comput., 17 (1996), pp. 1–15.
[15] L. N. Trefethen and M. Embree, *Spectra and pseudospectra*, Princeton University Press, Princeton, NJ, 2005.
[16] H. A. van der Vorst, *Iterative Krylov methods for large linear systems*, vol. 13 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2003.
[17] T. G. Wright and L. N. Trefethen, *Large-scale computation of pseudospectra using ARPACK and eigs*, SIAM J. Sci. Comput., 23 (2001), pp. 591–605.