Bounds on Eigenpairs in the Interior of Krylov Spaces

A D Kennedy

Monday, 1 August 2016

Introduction

- Krylov space approximations may used to compute the eigenpairs of a (large sparse) symmetric matrix A.
- In many applications, including lattice QCD, such matrices are approximations to an underlying differential operator.
- There are two aspects to understanding how effective Krylov space algorithms, such as the Lánczos algorithm:
  - How rapidly the eigenpairs of the projection H of A onto the Krylov space approach those of A as the dimension of the Krylov space increases, and
  - The numerical stability of the algorithm used to compute the eigenpairs of H.
- Our principal goal is to obtain a priori bounds (ones that essentially only depend on the spectrum of A) that address the first question for eigenvalues that lie in the interior of the spectrum.
- Our approach is a generalization of the textbook methods of Kaniel, Paige, and Saad that give such bounds for extremal eigenpairs.

Previous Work

- A significant application in lattice QCD is the computation of the eigenpairs in the middle of the spectrum of the Hermitian Dirac operator $H = \gamma_5 \not{D}$. [Edwards, Heller, and Narayanan, 1999]
- This made use of the Ritz algorithm, which minimizes the Rayleigh–Ritz functional $\rho\left(A^2, x\right) = \left(x, A^2 x\right)$ with respect to x where x is a unit vector. [Bunk, Jansen, Lüscher, and Simma, 1994; Kalkreuter and Simma, 1996]
- Clearly the interior eigenpairs of A correspond to the extremal eigenpairs of $(A - \Lambda)^2$ for some constant shift Λ.
- A problem is that the eigenspaces corresponding to eigenvalues $\Lambda \pm \lambda$ of A become degenerate, and it may be difficult to resolve them.

Basic Idea of Our Approach

- Our analysis is based upon the following trivial observation: the Krylov space for A^2 of dimension d is a subspace of the Krylov space for A of dimension $2d - 1$.
- Hence if there is a good approximation for an extremal eigenpair in the former then it must also be a good approximation for an internal eigenpair in the latter.
- Such internal eigenpairs must therefore be well-approximated in the Krylov space for A.
- No resolution of degenerate eigenspaces is required.
- The squared matrix is only used for the theoretical analysis.
- In practice one just constructs the Krylov space for A in the usual way, using the Lánczos algorithm or some variant (such as Partial Selective Orthogonalization, which we shall briefly descibe if time permits).

Krylov Spaces

We shall now give a brief introduction to Krylov spaces and Kaniel–Paige theory. For experts this can be taken to be a verbose summary of our notation.

- Let $A : \mathbb{R}^N \to \mathbb{R}^N$ be a linear operator (everything can be generalized to $\mathbb{C}^N$, of course).
- A is symmetric with respect to the inner product (x, y) of vectors $x, y \in \mathbb{R}^N$, i.e., $(x, Ay) = (Ax, y)$.
- We shall denote the norm of x by $\|x\| = \sqrt{(x,x)}$.
- The spectrum $\sigma(A)$ is the set of all the distinct eigenvalues of A, which we assume are labelled in increasing order

$$\boxed{\lambda_1 < \lambda_2 < \lambda_3 < \cdots < \lambda_{N'}.}$$

- The corresponding eigenvectors in $\mathbb{R}^N$ with $\|z_j\| = 1$ will be written as $z_1, z_2, z_3, \ldots, z_{N'}$
- Given an initial vector $v \neq 0$ the Krylov space of dimension $d \leq N'$ is defined to be

$$\boxed{\mathcal{K}_d(A) = \operatorname{span}(v, Av, A^2v, \ldots, A^{d-1}v).}$$

- $N' \leq N$ is the smallest integer such that $\mathcal{K}_{N'+1}(A) = \mathcal{K}_{N'}(A)$.
- We require that $z_j \in \mathcal{K}_{N'}(A)$ (all the eigenspaces of $\mathcal{K}_d(A)$ are non-degenerate).

Ritz Pairs

- Let $P : \mathbb{R}^N \to \mathbb{R}^d$ be an $d \times N$ matrix whose rows are orthonormal basis vectors of $\mathcal{K}_d(A)$. Then
  - $P^T P : \mathbb{R}^N \to \mathbb{R}^N$ is an orthogonal projector onto the subspace of $\mathbb{R}^N$ corresponding to $\mathcal{K}_d(A)$,
  - $PP^T = 1$ is the $d \times d$ unit matrix on the Krylov space,
  - and $H = PAP^T$ is the $d \times d$ projection of A onto the Krylov space.
- The eigenpairs (θ_j, s_j) of H (with $s_j \in \mathbb{R}^d$ and $\|s_j\| = 1$),

$$(H - \theta_j)s_j = 0,$$

  are called Ritz pairs.
- We order the Ritz values such that

$$\boxed{\theta_1 < \theta_2 < \cdots < \theta_d.}$$

- The corresponding Ritz vectors in $\mathbb{R}^N$ are $y_1, y_2, \ldots, y_d$, that is $y_j = P^T s_j$ with

$$\boxed{P(A - \theta_j)y_j = 0.}$$

- Clearly $\theta_1 \geq \lambda_1$.

Rayleigh Quotient

- The Rayleigh quotient (or Ritz functional) is defined to be

$$\boxed{\rho(A,u) = \frac{(u,Au)}{(u,u)} = (\hat{u}, A\hat{u}).}$$

- Since $\sum_{j=1}^{d} y_j \otimes y_j^T = 1$ on $\mathcal{K}_d(A)$, we have

$$\rho(A,u) = \rho(A,\hat{u}) = \sum_{j=1}^{d} \theta_j |(\hat{u}, y_j)|^2 \geq \theta_1 \sum_{j=1}^{d} |(\hat{u}, y_j)|^2 = \theta_1 \|\hat{u}\|^2 = \theta_1,$$

  for any $u \in \mathcal{K}_d(A)$, thus $\rho(A,u) \geq \theta_1$.
- Hence we have

$$\boxed{\rho(A - \lambda_1, u) \geq \theta_1 - \lambda_1 \geq 0}$$

  for any $u \in \mathcal{K}_d(A)$.

Kaniel Theory

- Following Kaniel we introduce a vector $u \in \mathcal{K}_d(A)$ such that $\hat{u} = u/\|u\|$ is an approximation to the eigenvector z_1.
- Since u lies in a Krylov space $u = p(A)v$ where p is a polynomial of degree $d - 1$.
- We shall call such a polynomial for which

$$|p(\lambda_j)| \leq 1 \quad \forall\, j > 1$$

  and $|p(\lambda_1)|$ is large a Kaniel polynomial.
- The simplest choice is $p(\lambda) = T_{d-1}\big(\gamma(\lambda)\big)$ where $T_d(x) = \cos\big(d \cos^{-1}(x)\big)$ is a Chebyshev polynomial and γ is the linear mapping with $\gamma(\lambda_2) = 1$ and $\gamma(\lambda_{N'}) = -1$.

Bounds on Rayleigh Quotient

- Since $\sum_{j=1}^{N'} z_j \otimes z_j^T = 1$ on $\mathcal{K}_{N'}(A)$ and $u = p(A)v \in \mathcal{K}_d(A) \subseteq \mathcal{K}_{N'}(A)$, we have

$$u = \sum_{j=1}^{N'} z_j(z_j, p(A)v) = \sum_{j=1}^{N'} z_j(p(A)z_j, v) = \sum_{j=1}^{N'} z_j p(\lambda_j)(z_j, v),$$

$$\implies \quad \|u\|^2 = (u, u) = \sum_{j=1}^{N'} p(\lambda_j)^2 |(z_j, v)|^2 \geq p(\lambda_1)^2 |(z_1, v)|^2 = p(\lambda_1)^2 \|v_{\|}\|^2$$

where $v_{\|} = z_1(z_1, v)$ is the projection of v onto z_1.

- Moreover,

$$\rho(A - \lambda_1, u)\, \|u\|^2 = \sum_{j=1}^{N'} (\lambda_j - \lambda_1) p(\lambda_j)^2 |(v, z_j)|^2 = \sum_{j=2}^{N'} (\lambda_j - \lambda_1) p(\lambda_j)^2 |(v, z_j)|^2$$

$$\leq \sum_{j=2}^{N'} (\lambda_j - \lambda_1) |(v, z_j)|^2 \leq (\lambda_{N'} - \lambda_1) \|v_\perp\|^2$$

where $v_\perp = v - v_{\|}$.

Extremal Eigenvalue Bounds

- We have thus established that

$$\rho\left(A-\lambda_1,u\right) \leq \frac{\lambda_{N'}-\lambda_1}{p(\lambda_1)^2}\frac{\|v_\perp\|^2}{\|v_{||}\|^2} = \frac{\lambda_{N'}-\lambda_1}{p(\lambda_1)^2}\left(\tan\angle_{z_1.v}\right)^2 \tag{1}$$

  where $\angle_{z_1,v}$ is the angle between the initial vector v and the eigenvector z_1.
- We thus have the bounds

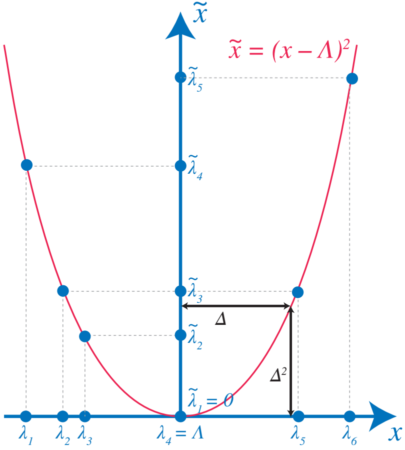
$$\boxed{0 \leq \theta_1 - \lambda_1 \leq \rho\left(A-\lambda_1,u\right) \leq \frac{\lambda_{N'}-\lambda_1}{p(\lambda_1)^2}\left(\tan\angle_{z_1,v}\right)^2.}$$

- The angle $\angle_{z_1,v}$ depends upon the initial vector v, and is thus not known *a priori*, however it does not depend upon the dimension d of the Krylov space (i.e., the number of Lánczos iterations).
- On the other hand the value of $p(\lambda_1)$ grows exponentially with d.

Folding

If z is an eigenvector of A belonging to eigenvalue λ then it must also be an eigenvector of $\tilde{A} = (A - \Lambda)^2$ with eigenvalue $\tilde{A}z = \tilde{\lambda}z = (\lambda - \Lambda)^2 z$.

1. The eigenvalues of A are plotted along the x-axis and those of $\tilde{A}$ along the $\tilde{x}$-axis.
2. Under the map $A \mapsto (A - \Lambda)^2$ the eigenvalues of A are lifted onto the red parabola.
3. Here the shift $\Lambda = \lambda_4$, so $\lambda_4 \mapsto \tilde{\lambda}_1 = 0$.
4. Since we label both sets of eigenvalues in increasing order the eigenvalue $\lambda_j \mapsto \tilde{\lambda}_{\pi_j}$: for example $\pi_2 = 3$ so $\lambda_2 \mapsto \tilde{\lambda}_3$.
5. Note that in this example the eigenvalues λ_2 and λ_5 of A both map to the same eigenvalue $\tilde{\lambda}_3$ of $\tilde{A}$.



Relationship between $\mathcal{K}_d(\tilde{A})$ and $\mathcal{K}_{2d}(A)$

If we take Λ to be the desired eigenvalue of A then

1. The vector $u = \tilde{p}(\tilde{A})v \in \mathcal{K}_d(\tilde{A})$ that approximates the eigenvector $\tilde{z}_1$ corresponding to the eigenvalue $\tilde{\lambda}_1 = 0$ in the folded spectrum also approximates $z_{\pi^{-1}(1)} = \tilde{z}_1$ of A with eigenvalue Λ.
2. Moreover $u \in \mathcal{K}_d(\tilde{A}) \subseteq \mathcal{K}_{2d-1}(A) \subseteq \mathcal{K}_{2d}(A)$, since any polynomial $\tilde{p}$ of degree $d-1$ in the variable $(x-\Lambda)^2$ is also a polynomial p of degree $2d-2$ in the variable x, where $\tilde{p}(x^2) = p(x)$.
3. It follows that $Au \in \mathcal{K}_{2d}(A)$.

Rayleigh Quotient for Kaniel vector u for $\tilde{A}$ in $\mathcal{K}_{2d}(A)$

- Since $(A-\Lambda)\hat{u} \in \mathcal{K}_{2d}(A)$ we may expand it in the complete orthonormal basis of Ritz vectors $y_1, y_2, \ldots, y_{2d}$ of $\mathcal{K}_{2d}(A)$ with corresponding Ritz values $\theta_1, \theta_2, \ldots, \theta_{2d}$:

$$\begin{aligned}
\rho\left(\tilde{A}, u\right) &= \rho\left((A-\Lambda)^2, \hat{u}\right) = \left(\hat{u}, (A-\Lambda)^2 \hat{u}\right) = \Big((A-\Lambda)\hat{u}, (A-\Lambda)\hat{u}\Big) \\
&= \sum_{i=1}^{2d} \sum_{j=1}^{2d} \Big((A-\Lambda)\hat{u}, y_i\Big)(y_i, y_j)\Big(y_j, (A-\Lambda)\hat{u}\Big) \\
&= \sum_{i=1}^{2d} \sum_{j=1}^{2d} \Big(\hat{u}, (A-\Lambda)y_i\Big)\delta_{ij}\Big((A-\Lambda)y_j, \hat{u}\Big) = \sum_{j-1}^{2d} (\theta_j - \Lambda)^2 |(y_j, \hat{u})|^2 \\
&= \sum_{j=1}^{2d} \tilde{\theta}_{\pi_j} |(y_j, \hat{u})|^2 \geq \tilde{\theta}_1 \|\hat{u}\|^2 = (\theta_{\pi^{-1}(1)} - \Lambda)^2.
\end{aligned}$$

- $\theta_{\pi^{-1}(1)}$ is the Ritz value closest to Λ, which may not be unique.

Interior Eigenvalue Bounds

- The (extremal) Kaniel bound for $u = \tilde{p}(\tilde{A})v$ is

$$\rho\left(\tilde{A}, u\right) \leq \frac{\tilde{\lambda}_{\tilde{N}'}}{\tilde{p}(0)^2} \left(\tan \angle_{\tilde{z}_1, v}\right)^2,$$

  since $\tilde{\lambda}_1 = 0$.

- Hence we obtain the bound on the (interior) Ritz value θ of A that is closest to Λ

$$\boxed{|\theta - \Lambda| \leq \sqrt{\rho\left(\tilde{A}, u\right)} \leq \left|\frac{(\tilde{\lambda} - \Lambda) \tan \angle_{\tilde{z}_1, v}}{\tilde{p}(0)}\right|}$$

  where $\tilde{\lambda} = \lambda_{\pi^{-1}(\tilde{N}')}$ is the eigenvalue of A furthest from Λ (i.e., $\tilde{\lambda}_{\tilde{N}'} = (\tilde{\lambda} - \Lambda)^2$), which again may not be unique.

- This bound behaves as $\tilde{p}(0)^{-1}$ for $2d - 1$ applications of A, as opposed to $p(\lambda_1)^{-2}$ for d applications of A in the extremal case; nevertheless it establishes exponential convergence in d.

Eigenvector Bounds

- It might seem desirable to have an *a priori* bound on the angle between the eigenvector z belonging to eigenvalue λ and the Ritz vector y belonging to the Ritz value θ closest to λ.
  - However, such a bound is often unsatisfactory, especially in the interior of the spectrum, because such a y and z may be completely unrelated.
  - A simple example of this phenomenon occurs when the spectrum $\sigma(A)$ is symmetric about zero but does not contain zero, and the initial vector v respects this symmetry. Under these conditions all odd-dimensional Krylov spaces $\mathcal{K}_{2d+1}(A)$ possess a vanishing Ritz value whose Ritz vector approximates no eigenvector of A.
  - Even for extremal eigenpairs such bounds can be very poor. For example, if $\lambda_1 \approx \lambda_2$ and $(v, z_1) \approx (v, z_2)$ then the eigenvectors z_1 and z_2 in $\mathcal{K}_{N'}(A)$ are orthogonal when $\lambda_2 - \lambda_1 = \varepsilon > 0$, but collapse onto a single eigenvector for $\varepsilon = 0$.
- We therefore derive an *a priori* bound on the angle $\angle_{z,\mathcal{Y}}$ between the eigenvector z and the subspace $\boxed{\mathcal{Y} = \mathrm{span}(y_1, y_2, \ldots, y_k)}$ corresponding to all the Ritz vectors with nearby Ritz values $|\theta_j - \lambda| < \Delta$ for $j = 1, 2, \ldots, k$ where Δ is a free parameter.

Eigensubspace Bounds

- We probably will not have time to go through the details of the calculation, but the result is that for the extremal eigenvector z_1 we obtain

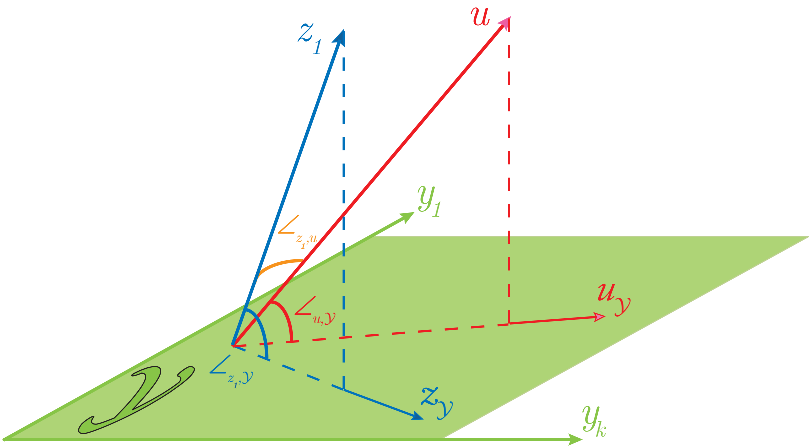
$$\boxed{|\angle_{z_1,\mathcal{Y}}| \leq \pi \left|\frac{\tan \angle_{z_1,v}}{p(\lambda_1)}\right| \left(\frac{1}{2} + \sqrt{\frac{\lambda_{N'} - \lambda_1}{\Delta}}\right)}$$

  where all the angles are assumed to be smaller than $\pi/2$.
- For an interior eigenvector $\tilde{z}_1$ the corresponding formula is very similar

$$\boxed{|\angle_{\tilde{z}_1,\mathcal{Y}}| \leq \pi \left|\frac{\tan \angle_{\tilde{z}_1,v}}{\tilde{p}(0)}\right| \left(\frac{1}{2} + \sqrt{\frac{\tilde{\lambda}_{\tilde{N}'}}{\Delta}}\right).}$$

- The geometric situation underlying the proof is illustrated in the following diagram.

Geometry


z_1
u
y_1
$\angle_{z_1,u}$
$\angle_{u,\mathcal{Y}}$
$u_{\mathcal{Y}}$
$\angle_{z_1,\mathcal{Y}}$
$\mathcal{Y}$
$z_{\mathcal{Y}}$
y_k

Proof of Extremal Eigenspace Bounds

- For any unit vectors x and y we have
$$\|x-y\|^2 = (x-y, x-y) = 2\Big(1-(x,y)\Big) = 2(1-\cos\angle_{x,y}) = (2\sin\tfrac{1}{2}\angle_{x,y})^2.$$
- The angle $\angle_{z_1,\mathcal{Y}}$ between the vector z_1 and the subspace $\mathcal{Y}$ is defined to be the angle between z_1 and the closest unit vector $z_\mathcal{Y} \in \mathcal{Y}$, so we have $\|z_1 - z_\mathcal{Y}\| = \min_{y\in\mathcal{Y}} \|z_1 - y\|$.
- Thus we have
$$\left|\sin\tfrac{1}{2}\angle_{z_1,\mathcal{Y}}\right| = \tfrac{1}{2}\|z_1 - z_\mathcal{Y}\|.$$
- Likewise, the nearest vector $u_\mathcal{Y} \in \mathcal{Y}$ to the Kaniel vector u satisfies
$$\left|\sin\tfrac{1}{2}\angle_{u,\mathcal{Y}}\right| = \tfrac{1}{2}\|\hat{u} - u_\mathcal{Y}\| = \tfrac{1}{2}\min_{y\in\mathcal{Y}}\|\hat{u} - y\|.$$
- As $u_\mathcal{Y} \in \mathcal{Y}$ we must have
$$\left|\sin\tfrac{1}{2}\angle_{z_1,\mathcal{Y}}\right| = \tfrac{1}{2}\min_{y\in\mathcal{Y}}\|z_1 - y\| \le \tfrac{1}{2}\|z_1 - u_\mathcal{Y}\|,$$
and since $\|z_1 - u_\mathcal{Y}\| = \|z_1 - u + u - u_\mathcal{Y}\| \le \|z_1 - u\| + \|u - u_\mathcal{Y}\|$ by the triangle inequality, we obtain
$$\left|\sin\tfrac{1}{2}\angle_{z_1,\mathcal{Y}}\right| \le \left|\sin\tfrac{1}{2}\angle_{z_1,u}\right| + \left|\sin\tfrac{1}{2}\angle_{u,\mathcal{Y}}\right|. \tag{2}$$

Extremal Bound on $\angle_{z_1,u}$

- Consider the ratio of the magnitude of the component $u_\perp$ of u perpendicular to z_1 to that of the component $u_{||}$ parallel to it,

$$
\begin{aligned}
(\tan\angle_{z_1,u})^2 &= \frac{\|u_\perp\|^2}{\|u_{||}\|^2} = \frac{\|u-(u,z_1)z_1\|^2}{\|(u,z_1)z_1\|^2} = \frac{\sum_{i=2}^{N'}|(u,z_i)|^2}{|(u,z_1)|^2} \\
&= \frac{\sum_{i=2}^{N'}|(p(A)v,z_i)|^2}{|(p(A)v,z_1)|^2} = \frac{\sum_{i=2}^{N'}|(v,z_i)p(\lambda_i)|^2}{|(v,z_1)p(\lambda_1)|^2} \leq \frac{\sum_{i=2}^{N'}|(v,z_i)|^2}{|(v,z_1)|^2 p(\lambda_1)^2} \\
&= \frac{\|z_\perp\|^2}{\|z_{||}\|^2 p(\lambda_1)^2} = \frac{(\tan\angle_{z_1,v})^2}{p(\lambda_1)^2}.
\end{aligned}
$$

- Therefore

$$
\boxed{|\tan\angle_{z_1,u}| \leq \left|\frac{\tan\angle_{z_1,v}}{p(\lambda_1)}\right|.} \tag{3}
$$

Extremal Bound on $\angle_{u,\mathcal{Y}}$

- Choose k such that $\theta_k - \lambda_1 < \Delta < \theta_{k+1} - \lambda_1$.
- The Rayleigh quotient is

$$\rho(A - \lambda_1, u) = \left(\hat{u}, (A - \lambda_1)\hat{u}\right) = \sum_{j=1}^{d} (\theta_j - \lambda_1) |(\hat{u}, y_j)|^2$$

$$\geq \sum_{j=k+1}^{d} (\theta_j - \lambda_1) |(\hat{u}, y_j)|^2 \geq (\theta_{k+1} - \lambda_1) \left(1 - \sum_{j=1}^{k} |(\hat{u}, y_j)|^2\right).$$

- Since $(\cos\angle_{u,\mathcal{Y}})^2 = |(\hat{u}, \hat{u}_{\mathcal{Y}})|^2 = \sum_{j=1}^{k} |(\hat{u}, y_j)|^2$ we have

$$\rho(A - \lambda_1, u) > \Delta (\sin\angle_{u,\mathcal{Y}})^2$$

.

- Hence, making use of equation (1),

$$\boxed{(\sin\angle_{u,\mathcal{Y}})^2 < \frac{\rho(A - \lambda_1, u)}{\Delta} \leq \frac{\lambda_{N'} - \lambda_1}{\Delta} \left|\frac{\tan\angle_{z_1,v}}{p(\lambda_1)}\right|^2.} \qquad (4)$$

Trigonometric Inequalities

We will make use of the following trivial inequalities

1. For $0 \leq x \leq \frac{\pi}{2}$ we have
$$\tan x = \frac{2\sin\frac{x}{2}\cos\frac{x}{2}}{\left(\cos\frac{x}{2}\right)^2 - \left(\sin\frac{x}{2}\right)^2} = \frac{2\tan\frac{x}{2}}{1-\left(\tan\frac{x}{2}\right)^2} \geq 2\tan\frac{x}{2} \geq 2\sin\frac{x}{2}.$$
2. Moreover, for $0 \leq x \leq \frac{2\pi}{3}$ we have
$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} \quad \text{and} \quad \tfrac{1}{2} \leq \cos\frac{x}{2} \leq 1 \implies \sin\frac{x}{2} \leq \sin x \leq 2\sin\frac{x}{2}.$$
3. Furthermore, for $0 \leq x \leq \frac{\pi}{2}$ we have $\frac{2x}{\pi} \leq \sin x \leq x$.

Bound on $\angle_{z_1,\mathcal{Y}}$

- Combining the bounds of equations (2), (3), and (4) we have

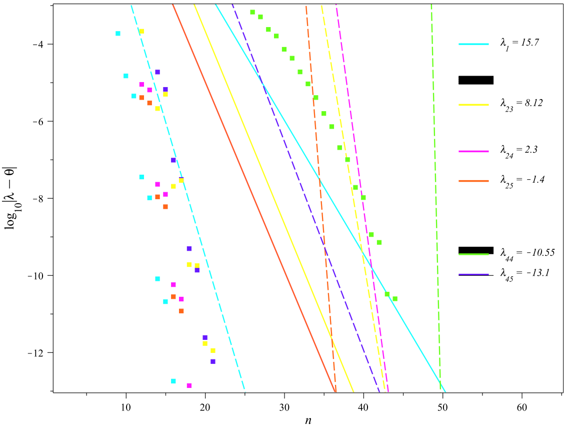
$$\left|\angle_{z_1,\mathcal{Y}}\right| \leq \frac{\pi}{2}\left|\sin\angle_{z_1,\mathcal{Y}}\right| \leq \pi\left|\sin\tfrac{1}{2}\angle_{z_1,\mathcal{Y}}\right| \leq \pi\left(\left|\sin\tfrac{1}{2}\angle_{z_1,u}\right| + \left|\sin\tfrac{1}{2}\angle_{u,\mathcal{Y}}\right|\right)$$
$$\leq \pi\left(\left|\tfrac{1}{2}\tan\angle_{z_1,u}\right| + \left|\sin\angle_{u,\mathcal{Y}}\right|\right) \leq \pi\left|\frac{\tan\angle_{z_1,v}}{p(\lambda_1)}\right|\left(\frac{1}{2} + \sqrt{\frac{\lambda_{N'}-\lambda_1}{\Delta}}\right)$$

  where all the angles are assumed to be smaller than $\pi/2$.
- The proof of the bounds given earlier for interior eigenvectors is quite similar.

Numerical Example

Convergence of Ritz values for matrix A whose spectrum is shown on the right.

1. The plot shows the logarithm of the magnitude of the difference between each Ritz value and the nearest eigenvalue of A as a function of the dimension n of the Krylov space.
2. The initial vector for the Krylov spaces was taken to have exactly equal overlap with each eigenvector.
3. The black "bands" in the spectrum are equally-spaced eigenvalues with a spacing of 0.05. The smallest eigenvalue is $\lambda_{46} = -13.2$.
4. The squares show the errors for the actual Ritz values, and their colour indicates the nearest eigenvalue.
5. The solid lines are the bounds from the shifted and squared matrix A' with shifts Λ chosen to give the smallest bound for an error of 10^{-8}, while the dashed lines are the original Kaniel–Paige–Saad bounds.
6. The shifts used were $\Lambda \to \infty$ for λ_1 and $\Lambda = 0.45$ for λ_{23}, λ_{24}, and λ_{25}.
7. The solid line for λ_{24} lies under that for λ_{25}: this is because they are equidistant from the shift value.
8. The Ritz values coincide with the eigenvalues for $n = \dim A = 46$



Lánczos Algorithm with Selective Orthogonalization

This algorithm is described in Parlett's book.

- The principal difficult in finding eigenpairs using the Lánczos algorithm is that eigenpairs appear multiple times due to rounding errors.
- Paige observed that eigenpairs reappear when they are well-approximated in the Krylov space: we call these good eigenpairs.
- Parlett and Scott introduced the Lánczos algorithm with Selective Orthogonalization (LANSO) to eliminate the spurious eigenpairs by explicitly reorthogonalizing the Lánczos vectors with respect to all previously found good eigenvectors.

Lánczos Algorithm with Partial Selective Orthogonalization

The details of this algorithm and its implementation were presented by Chris Johnson at the previous QCD&NA workshop, so we shall just provide the briefest of summaries here.

- As we are interested in eigenpairs with eigenvalues in a given small region of the spectrum we just reorthogonalize with respect to good eigenvectors whose eigenvalues lie in this region.
- We may easily find such good eigenvectors by using appropriate shifts in the QR algorithm to find candidate Ritz pairs (θ_j, s_j), and selecting those whose final components are smaller than $\sqrt{\epsilon}$ where ϵ is the unit of least precision (ULP) in floating point arithmetic (i.e., ϵ is the smallest floating point number such that $1 \oplus \epsilon > 1$.

Remarks and Conclusions

- In principle Krylov space methods only allow us to compute a single eigenvector belonging to a degenerate eigenvalue.
- In practice they give all the eigenvectors with the correct multiplicity thanks to rounding errors.
- We have shown why eigenpairs in low density regions of the spectrum are well-approximated by Ritz vectors in Krylov spaces.
- This result extends the well-known results for extremal eigenpairs.
- We have shown that the subspace of Ritz vectors with Ritz values close to an eigenvalue contains a good approximation to the corresponding eigenvector.
- Such subspace bounds are better behaved that eigenvector bounds even for extermal eigenpairs.
- This work was carried out in collaboration with Chris Johnson (EPCC).