Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe

http://www.tu-harburg.de/ins/veranstaltungen/gamm08/index.html

NOVELL: Support | Entertainment | News | Internet Search | Reference | Maps and Directions | Shopping | People and Companies

Technische Universität Hamburg-Harburg
Institut für Numerische Simulation

TUHH > Institut für Numerische Simulation > GAMM 2008

- Home
- Registration
- Program
- Abstracts
- Participants
- How to find us
- Accommodation
- Organizers

Supported by

- Bundesministerium für Bildung und Forschung
- GAMM

History

- Harrachov 2007
- Düsseldorf 2006
- Dresden 2005
- Hagen 2004
- Braunschweig 2003
- Bielefeld 2002
- Berlin 2001

Welcome to the GAMM Workshop

Applied and Numerical Linear Algebra

with special emphasis on

Regularization of Ill-posed Problems

Date:

September 11-12, 2008
Technische Universität Hamburg-Harburg, Germany

Invited speakers (confirmed):

Lars Eldén (Linköping, Sweden)

Per Christian Hansen (Lyngby, Denmark)

Marielba Rojas (Lyngby, Denmark)

Fiorella Sgallari (Bologna,Italy)

last modified on: 2008-04-14, 15:55

Quasi-Minimal Residual Eigenpairs

Jens-Peter M. Zemke
zemke@tu-harburg.de

Institut für Numerische Simulation
Technische Universität Hamburg-Harburg

10.09.2008
9.50 am – 10.15 am

IWASEP 7
June 9-12, 2008
Dubrovnik, Croatia

Outline

Krylov decompositions

We consider a given Krylov decomposition

$$AQ_k = Q_{k+1}\underline{C}_k = Q_kC_k + q_{k+1}c_{k+1,k}e_k^T. \tag{1}$$

Krylov decompositions

We consider a given Krylov decomposition

$$AQ_k = Q_{k+1}\underline{C}_k = Q_k C_k + q_{k+1} c_{k+1,k} e_k^T. \qquad (1)$$

We suppose that

$$A \in \mathbb{C}^{(n,n)}$$ is a general square matrix,

$$Q_{k+1} = \begin{pmatrix} Q_k & q_{k+1} \end{pmatrix} \in \mathbb{C}^{(n,k+1)}$$ is a matrix of basis vectors,

$$\underline{C}_k = \begin{pmatrix} C_k \\ c_{k+1,k} e_k^T \end{pmatrix} \in \mathbb{C}^{(k+1,k)}$$ is unreduced extended Hessenberg.

Krylov decompositions

We consider a given Krylov decomposition

$$AQ_k = Q_{k+1}\underline{C}_k = Q_kC_k + q_{k+1}c_{k+1,k}e_k^T. \tag{1}$$

We suppose that

$$A \in \mathbb{C}^{(n,n)}$$ is a general square matrix,

$$Q_{k+1} = \begin{pmatrix} Q_k & q_{k+1} \end{pmatrix} \in \mathbb{C}^{(n,k+1)}$$ is a matrix of basis vectors,

$$\underline{C}_k = \begin{pmatrix} C_k \\ c_{k+1,k}e_k^T \end{pmatrix} \in \mathbb{C}^{(k+1,k)}$$ is unreduced extended Hessenberg.

We do not consider perturbations.

Krylov decompositions

We consider a given Krylov decomposition

$$AQ_k = Q_{k+1}\underline{C}_k = Q_k C_k + q_{k+1}c_{k+1,k}e_k^T. \tag{1}$$

We suppose that

$$A \in \mathbb{C}^{(n,n)} \quad \text{is a general square matrix,}$$

$$Q_{k+1} = \begin{pmatrix} Q_k & q_{k+1} \end{pmatrix} \in \mathbb{C}^{(n,k+1)} \quad \text{is a matrix of basis vectors,}$$

$$\underline{C}_k = \begin{pmatrix} C_k \\ c_{k+1,k}e_k^T \end{pmatrix} \in \mathbb{C}^{(k+1,k)} \quad \text{is unreduced extended Hessenberg.}$$

We do not consider perturbations. We remark that important parts of the results carry over to general rectangular approximations $\underline{C}_k$ of A which not necessarily have to be Hessenberg.

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z, y = Q_k v} \frac{\|zy - Ay\|}{\|y\|} \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|zQ_k v - AQ_k v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|zQ_k v - AQ_k v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|zQ_k v - Q_{k+1}\underline{C}_k v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|zQ_k v - Q_{k+1}\underline{C}_k v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|zQ_{k+1}\underline{I}_k v - Q_{k+1}\underline{C}_k v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|zQ_{k+1}\underline{I}_k v - Q_{k+1}\underline{C}_k v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|Q_{k+1}(z\underline{I}_k - \underline{C}_k)v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} = \min_{z,v} \frac{\|Q_{k+1}(z\underline{I}_k - \underline{C}_k)v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \|Q_{k+1}\| \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|Q_k v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \|Q_{k+1}\| \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|Q_k v\|}. \qquad (2)$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \frac{\|Q_{k+1}\|}{\sigma_{\min}(Q_k)} \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \frac{\|Q_{k+1}\|}{\sigma_{\min}(Q_k)} \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \frac{\|Q_{k+1}\|}{\sigma_{\min}(Q_{k+1})} \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \frac{\|Q_{k+1}\|}{\sigma_{\min}(Q_{k+1})} \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \frac{\sigma_{\max}(Q_{k+1})}{\sigma_{\min}(Q_{k+1})} \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \frac{\sigma_{\max}(Q_{k+1})}{\sigma_{\min}(Q_{k+1})} \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \kappa(Q_{k+1}) \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \kappa(Q_{k+1}) \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

We always suppose that the columns of Q_{k+1} have been scaled to unit length.

QMR eigenpairs

We proceed similar to the QMR approach often applied to linear systems,

$$\min_{z,y=Q_k v} \frac{\|zy - Ay\|}{\|y\|} \leqslant \kappa(Q_{k+1}) \cdot \min_{z,v} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}. \tag{2}$$

We always suppose that the columns of Q_{k+1} have been scaled to unit length.

Definition (QMR eigenpair)

The pair $(\grave{\theta}, \grave{y} = Q_k \grave{v})$ is a QMR eigenpair, iff

$$\frac{\|(\grave{\theta}\underline{I}_k - \underline{C}_k)\grave{v}\|}{\|\grave{v}\|} = \operatorname*{min\,loc}_{z\in\mathbb{C}, v\in\mathbb{C}^k, \|v\|=1} \frac{\|(z\underline{I}_k - \underline{C}_k)v\|}{\|v\|}, \tag{3}$$

where “min loc” denotes a (not necessarily strict) local minimum.

QMR eigenpairs: SVD characterization

We denote the SVD of ${}^{z}\underline{C}_k \equiv z\underline{I}_k - \underline{C}_k$ by $U(z)\Sigma(z)V(z)^H = U\,\Sigma(z)V^H$.

QMR eigenpairs: SVD characterization

We denote the SVD of ${}^z\underline{C}_k \equiv z\underline{I}_k - \underline{C}_k$ by $U(z)\Sigma(z)V(z)^H = U\,\Sigma(z)V^H$.

Since for every $z \in \mathbb{C}$

$$\sigma_k(z) = \|\sigma_k(z)u_k\| = \frac{\|{}^z\underline{C}_k v_k\|}{\|v_k\|} = \min_v \frac{\|{}^z\underline{C}_k v\|}{\|v\|},$$

QMR eigenpairs: SVD characterization

We denote the SVD of ${}^{z}\underline{C}_k \equiv z\underline{I}_k - \underline{C}_k$ by $U(z)\Sigma(z)V(z)^H = U\,\Sigma(z)V^H$.

Since for every $z \in \mathbb{C}$

$$\sigma_k(z) = \|\sigma_k(z)u_k\| = \frac{\|{}^{z}\underline{C}_k v_k\|}{\|v_k\|} = \min_{v} \frac{\|{}^{z}\underline{C}_k v\|}{\|v\|},$$

the QMR eigenvalues can be characterized by

$$\grave{\theta} = \operatorname*{arg\,min\,loc}_{z\in\mathbb{C}} \sigma_k(z), \qquad (4)$$

QMR eigenpairs: SVD characterization

We denote the SVD of ${}^z\underline{C}_k \equiv z\underline{I}_k - \underline{C}_k$ by $U(z)\Sigma(z)V(z)^H = U\,\Sigma(z)V^H$.

Since for every $z \in \mathbb{C}$

$$\sigma_k(z) = \|\sigma_k(z)u_k\| = \frac{\|{}^z\underline{C}_k v_k\|}{\|v_k\|} = \min_v \frac{\|{}^z\underline{C}_k v\|}{\|v\|},$$

the QMR eigenvalues can be characterized by

$$\grave{\theta} = \arg\min_{z\in\mathbb{C}} \operatorname{loc} \sigma_k(z), \tag{4}$$

the QMR eigenvector can be chosen as a corresponding right singular vector,

$$\grave{v} = v_k(\grave{\theta}), \tag{5}$$

QMR eigenpairs: SVD characterization

We denote the SVD of ${}^{z}\underline{C}_k \equiv z\underline{I}_k - \underline{C}_k$ by $U(z)\Sigma(z)V(z)^H = U\,\Sigma(z)V^H$.

Since for every $z \in \mathbb{C}$

$$\sigma_k(z) = \|\sigma_k(z)u_k\| = \frac{\|{}^{z}\underline{C}_k v_k\|}{\|v_k\|} = \min_v \frac{\|{}^{z}\underline{C}_k v\|}{\|v\|},$$

the QMR eigenvalues can be characterized by

$$\grave{\theta} = \operatorname*{arg\,min\,loc}_{z\in\mathbb{C}} \sigma_k(z), \tag{4}$$

the QMR eigenvector can be chosen as a corresponding right singular vector,

$$\grave{v} = v_k(\grave{\theta}), \tag{5}$$

the QMR eigenresidual is given by $\sigma_k(\grave{\theta})$.

QMR eigenpairs: SVD steepest descent

Simple singular values $\sigma(z)$ and corresponding singular vectors v_k, u_k of the complex matrices ${}^{z}\underline{C}_k = z\underline{I}_k - \underline{C}_k$ are real analytic (Sun, 1988),

$$\sigma(z+w) = \sigma(z) + \sigma_z(z)w + \sigma_{\bar{z}}(z)\overline{w} + O(|w|^2) \tag{6}$$

$$= \sigma(z) + 2\Re((u_k^H \underline{I}_k v_k)w) + O(|w|^2). \tag{7}$$

QMR eigenpairs: SVD steepest descent

Simple singular values $\sigma(z)$ and corresponding singular vectors v_k, u_k of the complex matrices ${}^z\underline{C}_k = z\underline{I}_k - \underline{C}_k$ are real analytic (Sun, 1988),

$$\sigma(z+w) = \sigma(z) + \sigma_z(z)w + \sigma_{\overline{z}}(z)\overline{w} + O(|w|^2) \tag{6}$$

$$= \sigma(z) + 2\Re((u_k^H \underline{I}_k v_k)w) + O(|w|^2). \tag{7}$$

We obtain steepest descent by subtracting the conjugate of the gradient $\sigma_z(z)$:

$$z_{\text{new}} = z - \alpha\, \overline{u_k^H \underline{I}_k v_k} = z - \alpha\, v_k^H \underline{I}_k^H u_k \tag{8}$$

$$= z - \frac{\alpha}{\sigma_k} v_k^H \underline{I}_k^H (z\underline{I}_k - \underline{C}_k) v_k = z - \frac{\alpha}{\sigma_k} v_k^H (zI_k - C_k) v_k. \tag{9}$$

QMR eigenpairs: SVD steepest descent

Simple singular values $\sigma(z)$ and corresponding singular vectors v_k, u_k of the complex matrices ${}^z\underline{C}_k = z\underline{I}_k - \underline{C}_k$ are real analytic (Sun, 1988),

$$\sigma(z+w) = \sigma(z) + \sigma_z(z)w + \sigma_{\bar{z}}(z)\overline{w} + O(|w|^2) \tag{6}$$

$$= \sigma(z) + 2\Re((u_k^H \underline{I}_k v_k)w) + O(|w|^2). \tag{7}$$

We obtain steepest descent by subtracting the conjugate of the gradient $\sigma_z(z)$:

$$z_{\text{new}} = z - \alpha\,\overline{u_k^H \underline{I}_k v_k} = z - \alpha\, v_k^H \underline{I}_k^H u_k \tag{8}$$

$$= z - \frac{\alpha}{\sigma_k} v_k^H \underline{I}_k^H (z\underline{I}_k - \underline{C}_k) v_k = z - \frac{\alpha}{\sigma_k} v_k^H (zI_k - C_k) v_k. \tag{9}$$

We note that $\sigma_k(z)$ is the backward error of the approximate eigenvalue z. Setting $\alpha = \sigma_k$ yields alternating projections and is nearly optimal:

$$z_{\text{new}} = v_k^H C_k v_k. \tag{10}$$

QMR eigenpairs: SVD Newton

Steepest descent exhibits linear convergence. The real-analyticity of simple singular values can also be used to adopt Newton's method for stationary points.

QMR eigenpairs: SVD Newton

Steepest descent exhibits linear convergence. The real-analyticity of simple singular values can also be used to adopt Newton's method for stationary points.

Newton's method exhibits the usual locally quadratic convergence behavior, but in most cases for Newton's method good starting values have to be used, better than, say, the Ritz values.

QMR eigenpairs: SVD Newton

Steepest descent exhibits linear convergence. The real-analyticity of simple singular values can also be used to adopt Newton's method for stationary points.

Newton's method exhibits the usual locally quadratic convergence behavior, but in most cases for Newton's method good starting values have to be used, better than, say, the Ritz values.

In general Newton's method has problems with clustered and multiple singular values and if far from a solution, as the function to be optimized is almost linear far from stationary points.

QMR eigenpairs: SVD Newton

Steepest descent exhibits linear convergence. The real-analyticity of simple singular values can also be used to adopt Newton's method for stationary points.

Newton's method exhibits the usual locally quadratic convergence behavior, but in most cases for Newton's method good starting values have to be used, better than, say, the Ritz values.

In general Newton's method has problems with clustered and multiple singular values and if far from a solution, as the function to be optimized is almost linear far from stationary points.

Enhancement: Damped Newton's method or simply BFGS.

QMR eigenpairs: SVD Newton

Steepest descent exhibits linear convergence. The real-analyticity of simple singular values can also be used to adopt Newton's method for stationary points.

Newton's method exhibits the usual locally quadratic convergence behavior, but in most cases for Newton's method good starting values have to be used, better than, say, the Ritz values.

In general Newton's method has problems with clustered and multiple singular values and if far from a solution, as the function to be optimized is almost linear far from stationary points.

Enhancement: Damped Newton's method or simply BFGS.

Remark: Multiple singular values can not occur in the symmetric case due to the unreduced Hessenberg structure, but still may be pathologically close, compare with results by Lehmann and Wilkinson.

QMR eigenpairs: Grassmannian Optimization

Given an QMR eigenvector $\grave{v}$, we obtain $\grave{\theta}$ by the Rayleigh quotient with C_k, as

$$\grave{\theta} = \frac{\grave{v}^H C_k \grave{v}}{\grave{v}^H \grave{v}} = \operatorname*{arg\,min}_{z \in \mathbb{C}} \frac{\|(z\underline{I}_k - \underline{C}_k)\grave{v}\|}{\|\grave{v}\|}. \tag{11}$$

QMR eigenpairs: Grassmannian Optimization

Given an QMR eigenvector $\grave{v}$, we obtain $\grave{\theta}$ by the Rayleigh quotient with C_k, as

$$\grave{\theta} = \frac{\grave{v}^H C_k \grave{v}}{\grave{v}^H \grave{v}} = \operatorname*{arg\,min}_{z\in\mathbb{C}} \frac{\|(z\underline{I}_k - \underline{C}_k)\grave{v}\|}{\|\grave{v}\|}. \tag{11}$$

We could try to use only those z defined by the Rayleigh quotient, thus we set

$$z = z(v) = \frac{v^H C_k v}{v^H v}. \tag{12}$$

QMR eigenpairs: Grassmannian Optimization

Given an QMR eigenvector $\grave{v}$, we obtain $\grave{\theta}$ by the Rayleigh quotient with C_k, as

$$\grave{\theta} = \frac{\grave{v}^H C_k \grave{v}}{\grave{v}^H \grave{v}} = \arg\min_{z \in \mathbb{C}} \frac{\|(z\underline{I}_k - \underline{C}_k)\grave{v}\|}{\|\grave{v}\|}. \tag{11}$$

We could try to use only those z defined by the Rayleigh quotient, thus we set

$$z = z(v) = \frac{v^H C_k v}{v^H v}. \tag{12}$$

It can then be shown that the QMR eigenvectors give minima of the resulting real-analytic function

$$\lambda \,:\, G_1(\mathbb{C}^k) \to \mathbb{R}_{\geqslant 0},$$

$$\lambda \,:\, v \mapsto \lambda(v) = \frac{\|(z(v)\underline{I}_k - \underline{C}_k)v\|^2}{\|v\|^2} = \frac{v^H (\underline{C}_k)^H \underline{C}_k v}{v^H v} - \left|\frac{v^H C_k v}{v^H v}\right|^2. \tag{13}$$

QMR eigenpairs: Grassmannian Optimization

Given an QMR eigenvector $\grave{v}$, we obtain $\grave{\theta}$ by the Rayleigh quotient with C_k, as

$$\grave{\theta} = \frac{\grave{v}^H C_k \grave{v}}{\grave{v}^H \grave{v}} = \arg\min_{z \in \mathbb{C}} \frac{\|(z\underline{I}_k - \underline{C}_k)\grave{v}\|}{\|\grave{v}\|}. \qquad (11)$$

We could try to use only those z defined by the Rayleigh quotient, thus we set

$$z = z(v) = \frac{v^H C_k v}{v^H v}. \qquad (12)$$

It can then be shown that the QMR eigenvectors give minima of the resulting real-analytic function

$$\lambda \,:\, G_1(\mathbb{C}^k) \to \mathbb{R}_{\geqslant 0},$$
$$\lambda \,:\, v \mapsto \lambda(v) = \frac{\|(z(v)\underline{I}_k - \underline{C}_k)v\|^2}{\|v\|^2} = \frac{v^H(\underline{C}_k)^H \underline{C}_k v}{v^H v} - \left|\frac{v^H C_k v}{v^H v}\right|^2. \qquad (13)$$

The stationary points of real-analytic λ are always singular vectors.

QMR eigenpairs: Grassmannian Optimization

If the stationary point gives a minimum and the associated singular value is simple, then

QMR eigenpairs: Grassmannian Optimization

If the stationary point gives a minimum and the associated singular value is simple, then

- the stationary point $\grave{v} = v$ is an QMR eigenvector,

QMR eigenpairs: Grassmannian Optimization

If the stationary point gives a minimum and the associated singular value is simple, then

- the stationary point $\grave{v} = v$ is an QMR eigenvector,
- the associated QMR eigenvalue is characterized by $\grave{\theta} = \grave{v}^H C_k \grave{v}$,

QMR eigenpairs: Grassmannian Optimization

If the stationary point gives a minimum and the associated singular value is simple, then

- the stationary point $\grave{v} = v$ is an QMR eigenvector,
- the associated QMR eigenvalue is characterized by $\grave{\theta} = \grave{v}^H C_k \grave{v}$, and
- the QMR eigenresidual is given by $\sigma(\grave{\theta}) = \sqrt{\lambda(\grave{v})}$.

QMR eigenpairs: Grassmannian Optimization

If the stationary point gives a minimum and the associated singular value is simple, then

- the stationary point $\grave{v} = v$ is an QMR eigenvector,
- the associated QMR eigenvalue is characterized by $\grave{\theta} = \grave{v}^H C_k \grave{v}$, and
- the QMR eigenresidual is given by $\sigma(\grave{\theta}) = \sqrt{\lambda(\grave{v})}$.

We experimented with steepest descent and Newton's method for minimization of (the real-analytic) λ on the first (complex) Grassmannian in the framework of optimization on Riemannian manifolds (as recently developed by Smith; Edelman, Arias & Smith; Manton).

QMR eigenpairs: Grassmannian Optimization

If the stationary point gives a minimum and the associated singular value is simple, then

- the stationary point $\grave{v} = v$ is an QMR eigenvector,
- the associated QMR eigenvalue is characterized by $\grave{\theta} = \grave{v}^H C_k \grave{v}$, and
- the QMR eigenresidual is given by $\sigma(\grave{\theta}) = \sqrt{\lambda(\grave{v})}$.

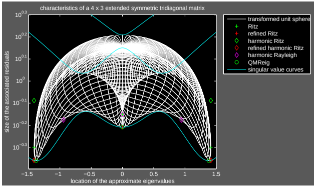
We experimented with steepest descent and Newton's method for minimization of (the real-analytic) λ on the first (complex) Grassmannian in the framework of optimization on Riemannian manifolds (as recently developed by Smith; Edelman, Arias & Smith; Manton).

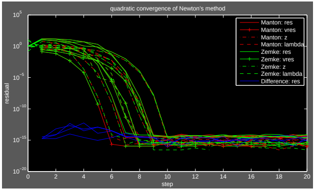
For Newton's method we have to compute the second covariant derivative, i.e., to use the Levi-Civita connection on the Grassmannian. This is simplified if using orthonormal frames, compare with the introductory textbook by Boothby.

QMR eigenpairs: Grassmannian Newton

```
for j = 1:convergence

  [Q,R] = qr(v); W = Q(:,2:k); v = Q(:,1);

  z = v'*Ck*v; zuCk = z*uIk-uCk;
  zuCkW = zuCk*W; zuCkv = zuCk*v;
  slambda = norm(zuCkv);
  y = zuCkW'*zuCkv;
  grad = 2*[real(y);imag(y)];
  res = norm(grad);

  A = zuCkW'*zuCkW; zCk = uIk'*zuCk;
  g1 = (zCk*W)'*v; r1 = real(g1); c1 = imag(g1);
  g2 = W'*(zCk*v); r2 = real(g2); c2 = imag(g2);
  outer1 = [r1+r2;c1+c2];
  outer2 = [c2-c1;r1-r2];
  Hesse = 2*[real(A) imag(A)';...
             imag(A) real(A)]-...
          2*slambda^2*I-...
          2*outer1*outer1'-2*outer2*outer2';
  ab = Hesse\grad;
  u = -W*(ab(1:k-1)+i*ab(k:2*k-2));
  normu = norm(u);
  v = v*cos(normu)+u*sin(normu)/normu;

end
```

(This is to convince you that the code is short enough to fit on one page.)





A graphical representation

We associate with every real or complex approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{v})$ a point (z, w) in the plane $\mathbb{R} \times \mathbb{R}$ or $\mathbb{C} \times \mathbb{R}$

A graphical representation

We associate with every real or complex approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{v})$ a point (z, w) in the plane $\mathbb{R} \times \mathbb{R}$ or $\mathbb{C} \times \mathbb{R}$:

$$z = \tilde{\theta}, \qquad w = \frac{\|(\tilde{\theta} \underline{I}_k - \underline{C}_k)\tilde{v}\|}{\|\tilde{v}\|}. \tag{14}$$

A graphical representation

We associate with every real or complex approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{v})$ a point (z, w) in the plane $\mathbb{R} \times \mathbb{R}$ or $\mathbb{C} \times \mathbb{R}$:

$$z = \tilde{\theta}, \qquad w = \frac{\|(\tilde{\theta}\underline{I}_k - \underline{C}_k)\tilde{v}\|}{\|\tilde{v}\|}. \tag{14}$$

The former gives the approximate eigenvalue, the latter gives the norm of the (quasi-)residual of the approximate eigenpair.

A graphical representation

We associate with every real or complex approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{v})$ a point (z, w) in the plane $\mathbb{R} \times \mathbb{R}$ or $\mathbb{C} \times \mathbb{R}$:

$$z = \tilde{\theta}, \qquad w = \frac{\|(\tilde{\theta} \underline{I}_k - \underline{C}_k)\tilde{v}\|}{\|\tilde{v}\|}. \tag{14}$$

The former gives the approximate eigenvalue, the latter gives the norm of the (quasi-)residual of the approximate eigenpair.

The norm of the residual of $(\tilde{\theta}, \tilde{y})$ gives the backward error, i.e.,

$$w = \min\left\{\|\Delta A\| \,:\, (A + \Delta A)\tilde{y} = \tilde{y}\,\tilde{\theta}\right\}. \tag{15}$$

A graphical representation

We associate with every real or complex approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{v})$ a point (z, w) in the plane $\mathbb{R} \times \mathbb{R}$ or $\mathbb{C} \times \mathbb{R}$:

$$z = \tilde{\theta}, \qquad w = \frac{\|(\tilde{\theta} \underline{I}_k - \underline{C}_k)\tilde{v}\|}{\|\tilde{v}\|}. \tag{14}$$

The former gives the approximate eigenvalue, the latter gives the norm of the (quasi-)residual of the approximate eigenpair.

The norm of the residual of $(\tilde{\theta}, \tilde{y})$ gives the backward error, i.e.,

$$w = \min \left\{ \|\Delta A\| \ : \ (A + \Delta A)\tilde{y} = \tilde{y}\,\tilde{\theta} \right\}. \tag{15}$$

Remark 1: Without additional knowledge a small backward error is the best we can achieve.

A graphical representation

We associate with every real or complex approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{v})$ a point (z, w) in the plane $\mathbb{R} \times \mathbb{R}$ or $\mathbb{C} \times \mathbb{R}$:

$$z = \tilde{\theta}, \qquad w = \frac{\|(\tilde{\theta}\underline{I}_k - \underline{C}_k)\tilde{v}\|}{\|\tilde{v}\|}. \tag{14}$$

The former gives the approximate eigenvalue, the latter gives the norm of the (quasi-)residual of the approximate eigenpair.

The norm of the residual of $(\tilde{\theta}, \tilde{y})$ gives the backward error, i.e.,

$$w = \min\left\{\|\Delta A\| \;:\; (A + \Delta A)\tilde{y} = \tilde{y}\,\tilde{\theta}\right\}. \tag{15}$$

Remark 1: Without additional knowledge a small backward error is the best we can achieve.

Remark 2: There exist “graphical” bounds for general and “Rayleigh” approximations.

A beautiful example

As an example we use

$$\underline{C}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

A beautiful example

As an example we use

$$\underline{C}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

Its Ritz values are given by

$$\theta_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \theta_2 = 0, \tag{17}$$

A beautiful example

As an example we use

$$\underline{C}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

Its Ritz values are given by

$$\theta_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \theta_2 = 0, \tag{17}$$

its harmonic Ritz values are given by

$$\underline{\theta}_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \underline{\theta}_2 = \infty, \tag{18}$$

A beautiful example

As an example we use

$$\underline{C}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

Its Ritz values are given by

$$\theta_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \theta_2 = 0, \tag{17}$$

its harmonic Ritz values are given by

$$\underline{\theta}_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \underline{\theta}_2 = \infty, \tag{18}$$

its ρ-values (Rayleigh quotients with harmonic Ritz vectors) are given by

$$\rho_{1,3} = \mp\sqrt{2} \cdot \frac{2}{3} \approx \mp 0.9428090, \quad \rho_2 = 0, \tag{19}$$

A beautiful example

As an example we use

$$\underline{C}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

Its Ritz values are given by

$$\theta_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \theta_2 = 0, \tag{17}$$

its harmonic Ritz values are given by

$$\underline{\theta}_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \underline{\theta}_2 = \infty, \tag{18}$$

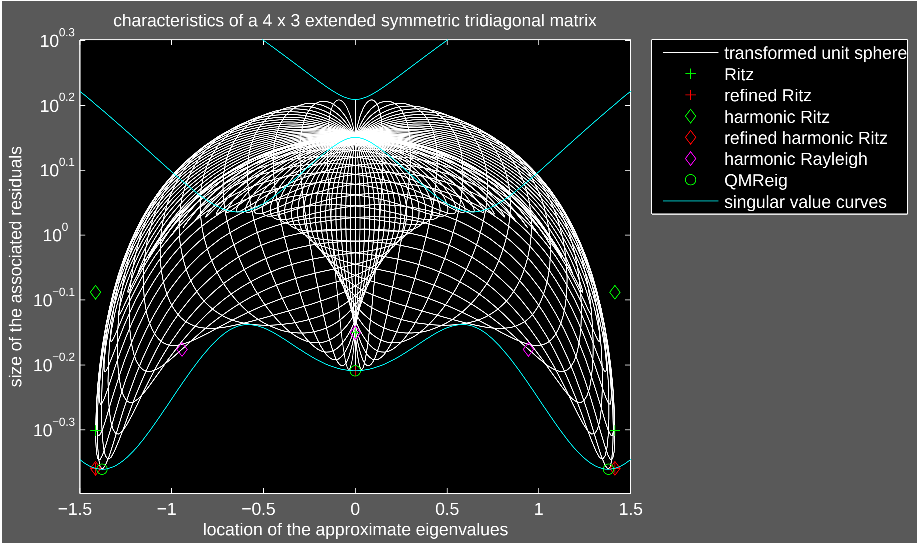
its ρ-values (Rayleigh quotients with harmonic Ritz vectors) are given by

$$\rho_{1,3} = \mp\sqrt{2} \cdot \frac{2}{3} \approx \mp 0.9428090, \quad \rho_2 = 0, \tag{19}$$

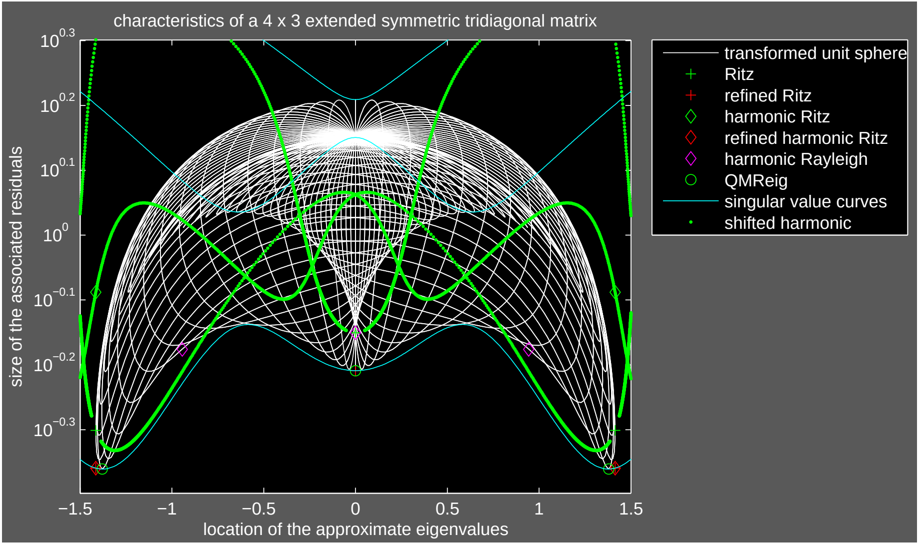
and its QMR eigenvalues are given by (where $y = 276081 + 21504\sqrt{2}i$)

$$\grave{\theta}_{1,3} = \mp\frac{\sqrt{2}}{16}\sqrt{113 + 2\Re\sqrt[3]{y}} \approx \mp 1.37898323557, \quad \grave{\theta}_2 = 0. \tag{20}$$

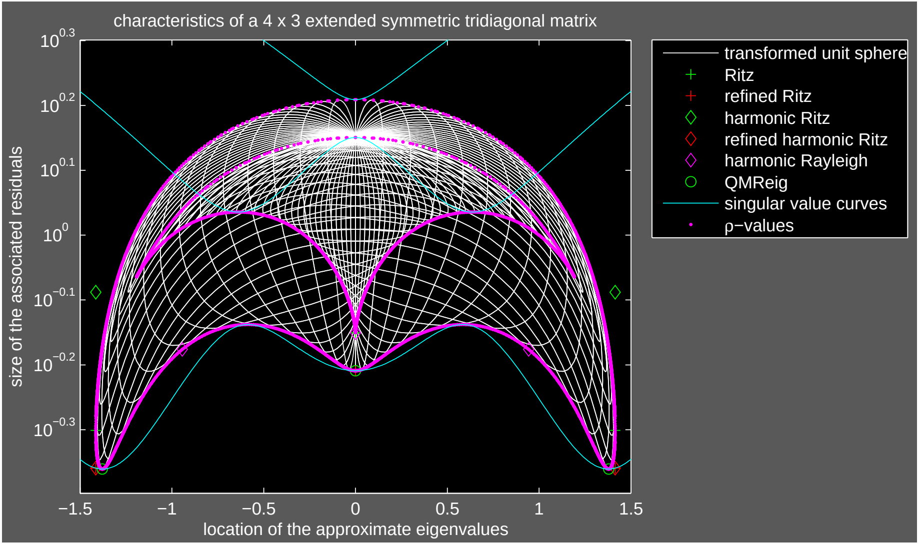
A beautiful example



A beautiful example



A beautiful example



Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.

Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.
- We have sketched several algorithms for the computation of QMR eigenpairs.

Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.
- We have sketched several algorithms for the computation of QMR eigenpairs.
- We indicated relations between known Krylov-subspace-based eigenpair extraction methods; those are based on the concept of QMR eigenpairs.

Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.
- We have sketched several algorithms for the computation of QMR eigenpairs.
- We indicated relations between known Krylov-subspace-based eigenpair extraction methods; those are based on the concept of QMR eigenpairs.
- We have introduced a powerful graphical concept which we think is capable to visualize and find important extraction related aspects.

Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.
- We have sketched several algorithms for the computation of QMR eigenpairs.
- We indicated relations between known Krylov-subspace-based eigenpair extraction methods; those are based on the concept of QMR eigenpairs.
- We have introduced a powerful graphical concept which we think is capable to visualize and find important extraction related aspects.
- We have weakened and generalized Lehmann's approaches for optimal eigenvalue inclusions.

Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.
- We have sketched several algorithms for the computation of QMR eigenpairs.
- We indicated relations between known Krylov-subspace-based eigenpair extraction methods; those are based on the concept of QMR eigenpairs.
- We have introduced a powerful graphical concept which we think is capable to visualize and find important extraction related aspects.
- We have weakened and generalized Lehmann's approaches for optimal eigenvalue inclusions.

Important open questions related to our concept include:

- What about the usefulness? Do we "need" QMR eigenpairs? Should they be "computed" or "otherwise approximated"?

Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.
- We have sketched several algorithms for the computation of QMR eigenpairs.
- We indicated relations between known Krylov-subspace-based eigenpair extraction methods; those are based on the concept of QMR eigenpairs.
- We have introduced a powerful graphical concept which we think is capable to visualize and find important extraction related aspects.
- We have weakened and generalized Lehmann's approaches for optimal eigenvalue inclusions.

Important open questions related to our concept include:

- What about the usefulness? Do we "need" QMR eigenpairs? Should they be "computed" or "otherwise approximated"?
- Are there applications?

Conclusion and Outview

- We have introduced the concept of QMR eigenpairs.
- We have sketched several algorithms for the computation of QMR eigenpairs.
- We indicated relations between known Krylov-subspace-based eigenpair extraction methods; those are based on the concept of QMR eigenpairs.
- We have introduced a powerful graphical concept which we think is capable to visualize and find important extraction related aspects.
- We have weakened and generalized Lehmann's approaches for optimal eigenvalue inclusions.

Important open questions related to our concept include:

- What about the usefulness? Do we "need" QMR eigenpairs? Should they be "computed" or "otherwise approximated"?
- Are there applications?
- Is there an "algebraic" characterization of all QMR eigenvalues?

Conclusion and Outview

The last item has been resolved partially, the QMR eigenvalues are the zeros of a polyanalytic polynomial of (in the generic case) total degree k^2, the other points are stationary points in both the SVD and Grassmannian descriptions.

Conclusion and Outview

The last item has been resolved partially, the QMR eigenvalues are the zeros of a polyanalytic polynomial of (in the generic case) total degree k^2, the other points are stationary points in both the SVD and Grassmannian descriptions.

This knowledge has been used together with some Computer Algebra (i.e., Maple) to compute the exact values of some QMR eigenvalues.

Conclusion and Outview

The last item has been resolved partially, the QMR eigenvalues are the zeros of a polyanalytic polynomial of (in the generic case) total degree k^2, the other points are stationary points in both the SVD and Grassmannian descriptions.

This knowledge has been used together with some Computer Algebra (i.e., Maple) to compute the exact values of some QMR eigenvalues.

There are cases of infinitely many QMR eigenpairs, namely, QMR eigenvalues on a circle, the zeros of $z\bar{z} - c^2 = 0$. Mostly, we obtain slightly less than k QMR eigenvalues. We need to know the number of zeros of a real algebraic variety.

Conclusion and Outview

The last item has been resolved partially, the QMR eigenvalues are the zeros of a polyanalytic polynomial of (in the generic case) total degree k^2, the other points are stationary points in both the SVD and Grassmannian descriptions.

This knowledge has been used together with some Computer Algebra (i.e., Maple) to compute the exact values of some QMR eigenvalues.

There are cases of infinitely many QMR eigenpairs, namely, QMR eigenvalues on a circle, the zeros of $z\bar{z} - c^2 = 0$. Mostly, we obtain slightly less than k QMR eigenvalues. We need to know the number of zeros of a real algebraic variety.

It remains an open question to come up with a “simple” description of these polyanalytic polynomials in terms of the matrices $\underline{C}_k$.

Conclusion and Outview

The last item has been resolved partially, the QMR eigenvalues are the zeros of a polyanalytic polynomial of (in the generic case) total degree k^2, the other points are stationary points in both the SVD and Grassmannian descriptions.

This knowledge has been used together with some Computer Algebra (i.e., Maple) to compute the exact values of some QMR eigenvalues.

There are cases of infinitely many QMR eigenpairs, namely, QMR eigenvalues on a circle, the zeros of $z\bar{z} - c^2 = 0$. Mostly, we obtain slightly less than k QMR eigenvalues. We need to know the number of zeros of a real algebraic variety.

It remains an open question to come up with a "simple" description of these polyanalytic polynomials in terms of the matrices $\underline{C}_k$.

Thank you for your attention!

A warning: loss of QMR eigenvalues; academic

An academic example is

$$\underline{C}_k = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{21}$$

A warning: loss of QMR eigenvalues; academic

An academic example is

$$\underline{C}_k = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{21}$$

In this special case (one Ritz value being zero), replacing "2" by an arbitrary number does not alter the Ritz, harmonic Ritz and ρ-values.

A warning: loss of QMR eigenvalues; academic

An academic example is

$$\underline{C}_k = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{21}$$

In this special case (one Ritz value being zero), replacing “2” by an arbitrary number does not alter the Ritz, harmonic Ritz and ρ-values.

Yet, this small change results in $\underline{C}_k$ having only two QMR eigenvalues, which are given by

$$\grave{\theta}_{1,2} = \mp\sqrt{\frac{11}{8}} \approx \mp 1.17260393996. \tag{22}$$

A warning: loss of QMR eigenvalues; academic

An academic example is

$$\underline{C}_k = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{21}$$

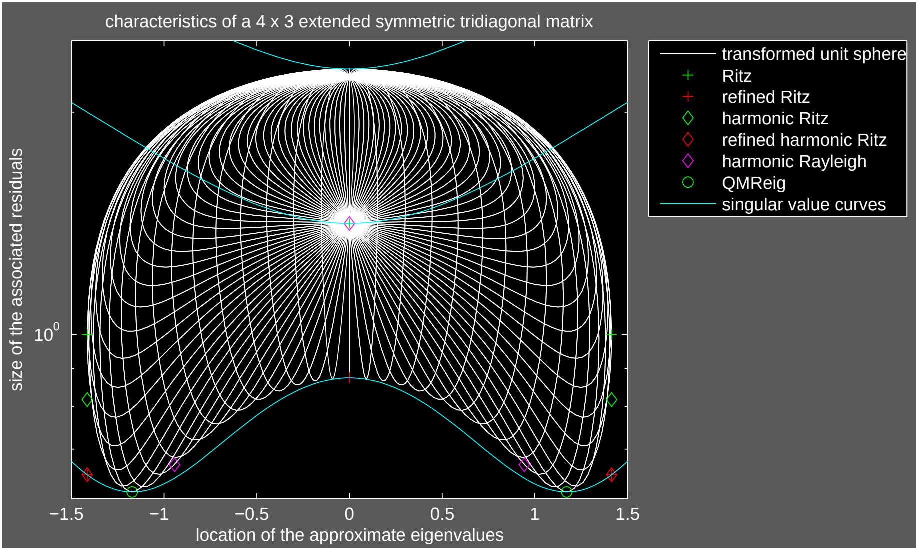
In this special case (one Ritz value being zero), replacing "2" by an arbitrary number does not alter the Ritz, harmonic Ritz and ρ-values.

Yet, this small change results in $\underline{C}_k$ having only two QMR eigenvalues, which are given by

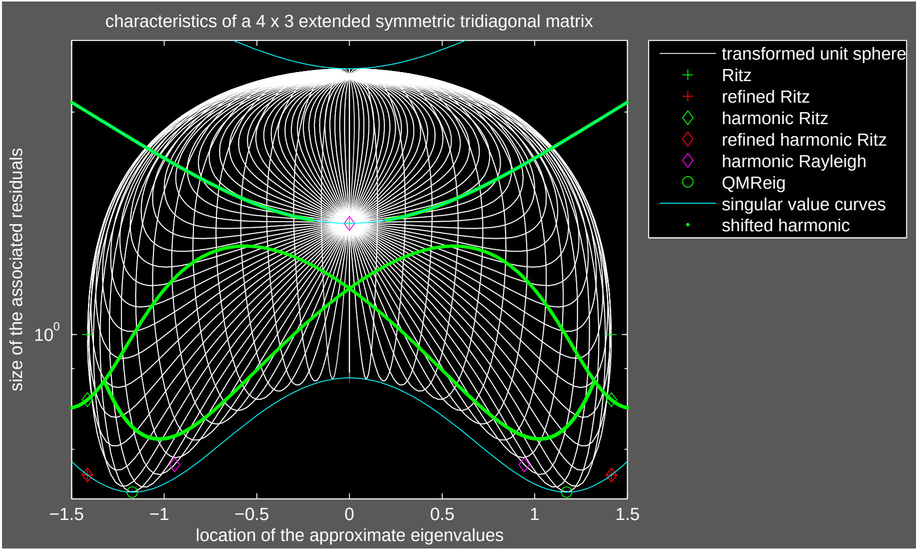
$$\grave{\theta}_{1,2} = \mp\sqrt{\frac{11}{8}} \approx \mp 1.17260393996. \tag{22}$$

At zero a stationary point exists which is a maximum of the smallest singular value curve and a saddle point of the transformed sphere.

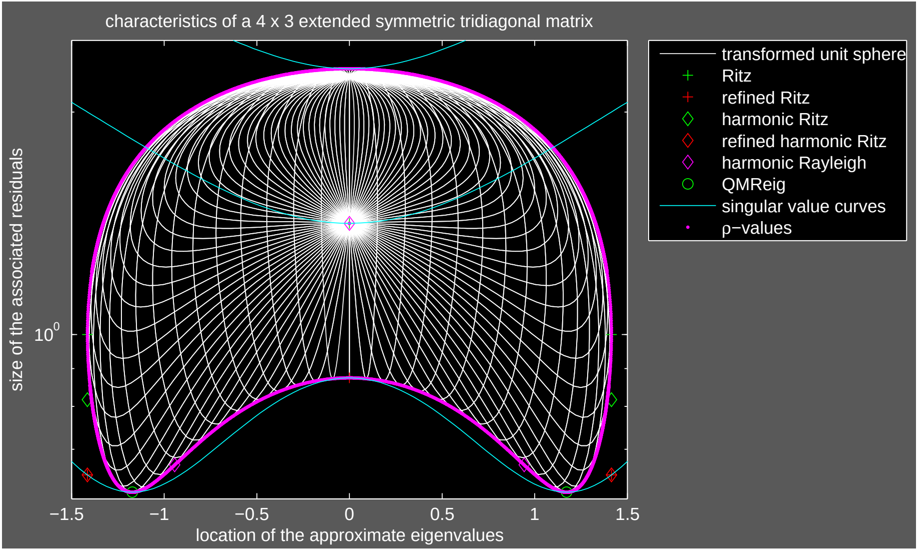
A warning: loss of QMR eigenvalues; academic



A warning: loss of QMR eigenvalues; academic



A warning: loss of QMR eigenvalues; academic



A warning: loss of QMR eigenvalue; realistic example

An interesting example is extended symmetric and generated using MATLAB's `randn` and `hess` functions and is approximately given by

$$\underline{C}_k \approx \begin{pmatrix} 0.46204801 & 1.75649255 & 0 \\ 1.75649255 & 0.23525002 & -0.70301190 \\ 0 & -0.70301190 & 1.90702012 \\ 0 & 0 & 1.21958322 \end{pmatrix}. \tag{23}$$

A warning: loss of QMR eigenvalue; realistic example

An interesting example is extended symmetric and generated using MATLAB's `randn` and `hess` functions and is approximately given by

$$\underline{C}_k \approx \begin{pmatrix} 0.46204801 & 1.75649255 & 0 \\ 1.75649255 & 0.23525002 & -0.70301190 \\ 0 & -0.70301190 & 1.90702012 \\ 0 & 0 & 1.21958322 \end{pmatrix}. \tag{23}$$

The computed Ritz, harmonic Ritz and ρ-values all differ.

A warning: loss of QMR eigenvalue; realistic example

An interesting example is extended symmetric and generated using MATLAB's `randn` and `hess` functions and is approximately given by

$$\underline{C}_k \approx \begin{pmatrix} 0.46204801 & 1.75649255 & 0 \\ 1.75649255 & 0.23525002 & -0.70301190 \\ 0 & -0.70301190 & 1.90702012 \\ 0 & 0 & 1.21958322 \end{pmatrix}. \qquad (23)$$

The computed Ritz, harmonic Ritz and ρ-values all differ. There are only two QMR eigenvalues.

A warning: loss of QMR eigenvalue; realistic example

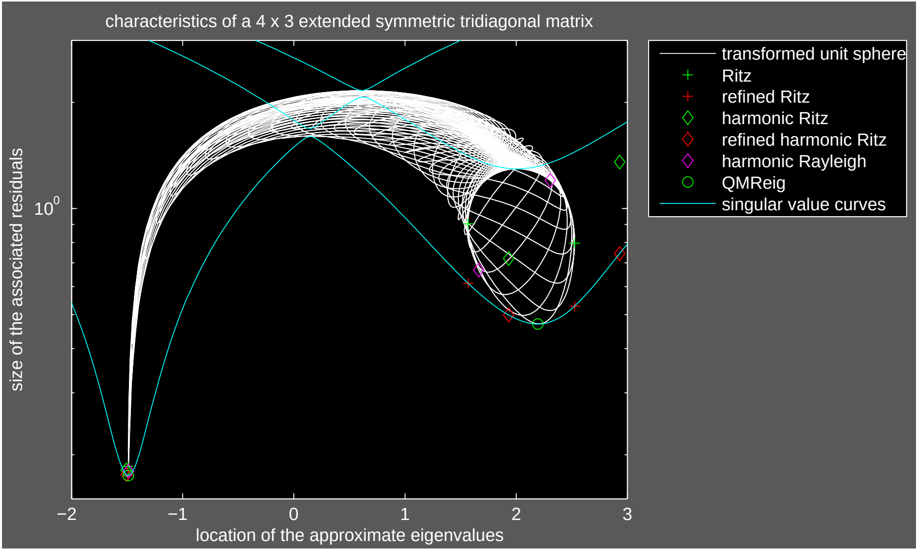
An interesting example is extended symmetric and generated using MATLAB's `randn` and `hess` functions and is approximately given by

$$\underline{C}_k \approx \begin{pmatrix} 0.46204801 & 1.75649255 & 0 \\ 1.75649255 & 0.23525002 & -0.70301190 \\ 0 & -0.70301190 & 1.90702012 \\ 0 & 0 & 1.21958322 \end{pmatrix}. \qquad (23)$$

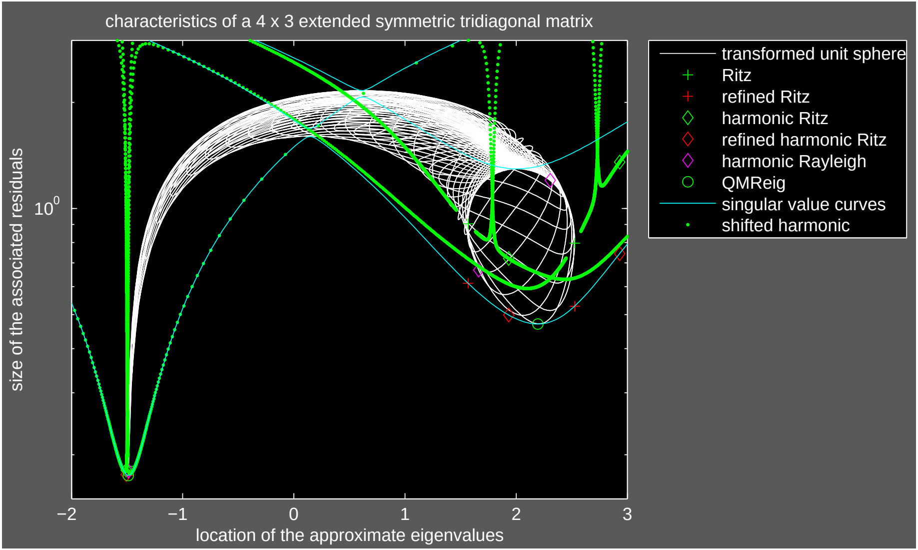
The computed Ritz, harmonic Ritz and ρ-values all differ. There are only two QMR eigenvalues. The smallest of all these and the norms of the eigenpair residuals (denoted by $n(\cdot,\cdot)$) are approximately given by

$$\begin{aligned} \theta_1 &\approx -1.490413407713866, & n(\theta_1, v_1) &\approx 0.185432088955641 7, \\ \underline{\theta}_1 &\approx -1.509143602001304, & n(\underline{\theta}_1, \underline{v}_1) &\approx 0.181039457164899 5, \\ \rho_1 &\approx -1.487425797938723, & n(\rho_1, \underline{v}_1) &\approx 0.179732084050847 2, \\ \grave{\theta}_1 &\approx -1.489367749116040, & n(\grave{\theta}_1, \grave{v}_1) &\approx 0.174658339265659 0. \end{aligned}$$

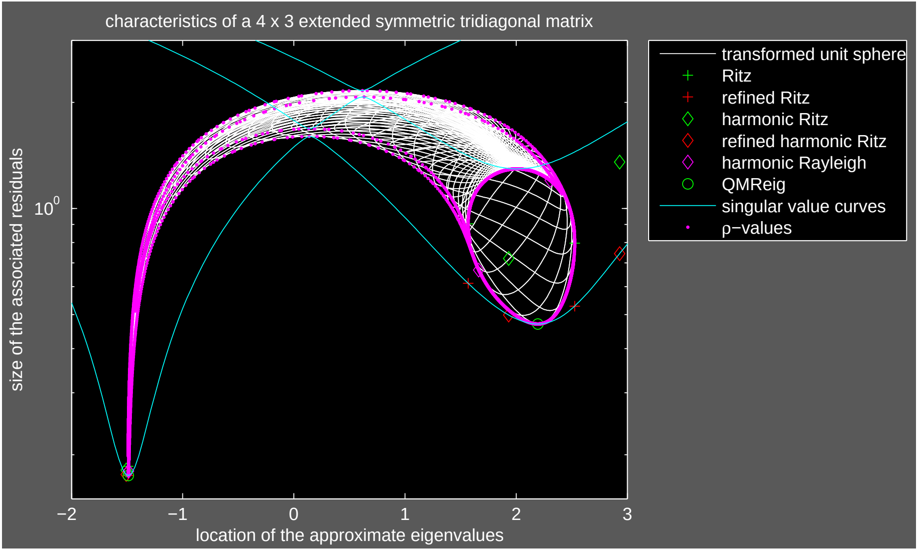
The ‘random’ example



The ‘random’ example



The ‘random’ example



Boundaries and ρ-values . . .

Why are the ρ-values on the "borders" of the transformed unit sphere?

Boundaries and ρ-values . . .

Why are the ρ-values on the "borders" of the transformed unit sphere?

In the symmetric case it is easy to characterize these "borders" and to prove that the vectors defining them are indeed harmonic Ritz vectors for two certain shifts. These shifts as well as the harmonic Ritz values are expressed using stationary points along straight lines in the graphical interpretation . . .

Boundaries and ρ-values . . .

Why are the ρ-values on the "borders" of the transformed unit sphere?

In the symmetric case it is easy to characterize these "borders" and to prove that the vectors defining them are indeed harmonic Ritz vectors for two certain shifts. These shifts as well as the harmonic Ritz values are expressed using stationary points along straight lines in the graphical interpretation . . .

Given the harmonic Ritz pair, it is even easier to find the direction along which the vector is a stationary point.

Boundaries and ρ-values ...

Why are the ρ-values on the "borders" of the transformed unit sphere?

In the symmetric case it is easy to characterize these "borders" and to prove that the vectors defining them are indeed harmonic Ritz vectors for two certain shifts. These shifts as well as the harmonic Ritz values are expressed using stationary points along straight lines in the graphical interpretation ...

Given the harmonic Ritz pair, it is even easier to find the direction along which the vector is a stationary point.

The QMR eigenvectors are harmonic Ritz vectors, the shifts are given by

$$\tau_{\pm} = \grave{\theta} \pm \sigma_k(\grave{\theta}), \tag{24}$$

in accordance with Lehmann's results for the symmetric case, see also van den Eshof's doctoral thesis (2003).

Boundaries and ρ-values . . .

Why are the ρ-values on the "borders" of the transformed unit sphere?

In the symmetric case it is easy to characterize these "borders" and to prove that the vectors defining them are indeed harmonic Ritz vectors for two certain shifts. These shifts as well as the harmonic Ritz values are expressed using stationary points along straight lines in the graphical interpretation . . .

Given the harmonic Ritz pair, it is even easier to find the direction along which the vector is a stationary point.

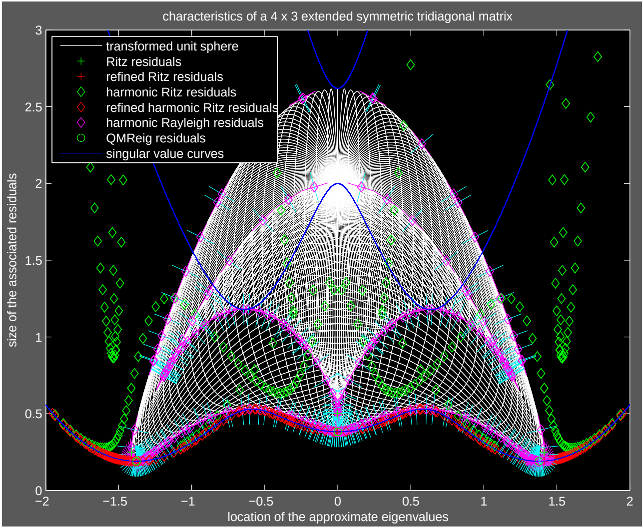
The QMR eigenvectors are harmonic Ritz vectors, the shifts are given by

$$\tau_{\pm} = \grave{\theta} \pm \sigma_k(\grave{\theta}), \tag{24}$$

in accordance with Lehmann's results for the symmetric case, see also van den Eshof's doctoral thesis (2003).

The Ritz vectors are harmonic Ritz vectors with shifts $\tau_{\pm} = \pm\infty$.

Boundaries and ρ-values . . .



Shifted harmonic Ritz values and the SVD

The harmonic Ritz values $\underline{\theta}$ are the eigenvalues of the inverse of a section of the pseudoinverse,

$$(\underline{C}_k^{\dagger}\underline{I}_k)^{-1}\underline{v} = \underline{v}\,\underline{\theta}$$

Shifted harmonic Ritz values and the SVD

The harmonic Ritz values $\underline{\theta}$ are the eigenvalues of the inverse of a section of the pseudoinverse,

$$(\underline{C}_k^{\dagger}\underline{I}_k)^{-1}\underline{v} = \underline{v}\,\underline{\theta}$$

The same is true for the shifted harmonic Ritz values, these are obtained as the same section of the pseudoinverse of the shifted rectangular matrices,

$$((\underline{C}_k - \tau\underline{I}_k)^{\dagger}\underline{I}_k)^{-1}\underline{v}(\tau) = \underline{v}(\tau)\,(\underline{\theta}(\tau) + \tau).$$

Shifted harmonic Ritz values and the SVD

The harmonic Ritz values $\underline{\theta}$ are the eigenvalues of the inverse of a section of the pseudoinverse,

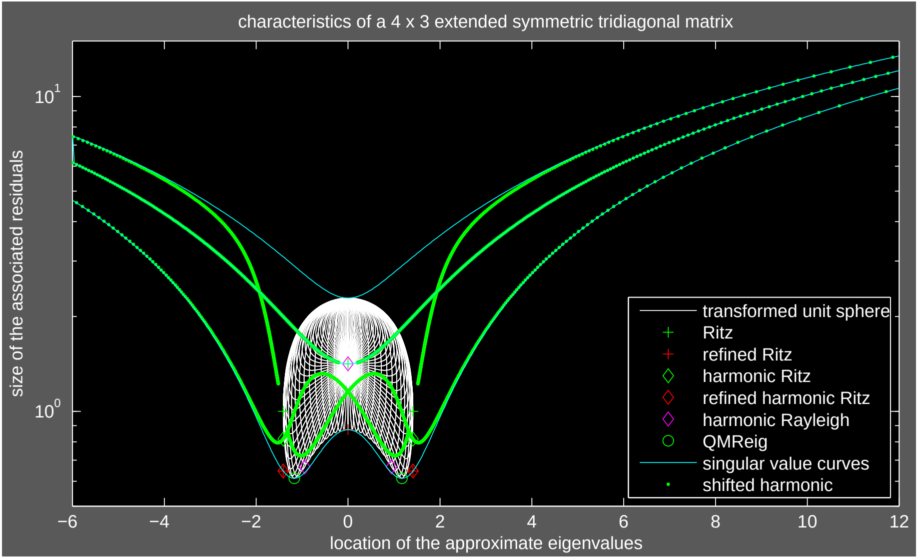
$$(\underline{C}_k^{\dagger}\underline{I}_k)^{-1}\underline{v} = \underline{v}\,\underline{\theta}$$

The same is true for the shifted harmonic Ritz values, these are obtained as the same section of the pseudoinverse of the shifted rectangular matrices,

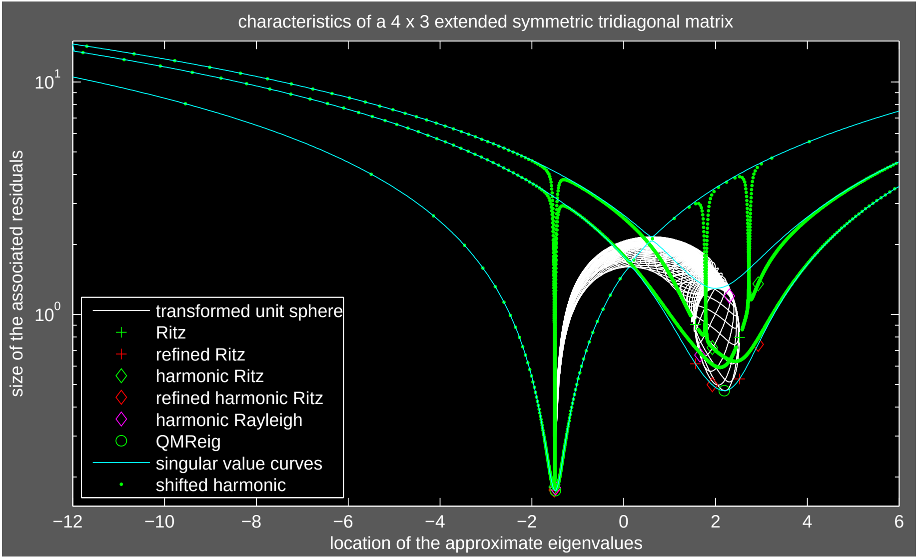
$$((\underline{C}_k - \tau\underline{I}_k)^{\dagger}\underline{I}_k)^{-1}\underline{v}(\tau) = \underline{v}(\tau)\,(\underline{\theta}(\tau) + \tau).$$

As the harmonic Ritz values are not shift-invariant (in contrast to Ritz and QMR eigenvalues), and by interlacing of singular values and the connection of the pseudoinverse to the singular value decomposition we might expect to see relations between shifted harmonic Ritz and the SVD in the pictures.

Shifted harmonic Ritz values and the SVD



Shifted harmonic Ritz values and the SVD



A Jordan block: infinitely many QMR eigenvalues

A startling example used already by Eising in context of the distance to uncontrollability (given in terms of the the best QMR eigenpair) is

$$\underline{C}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{25}$$

A Jordan block: infinitely many QMR eigenvalues

A startling example used already by Eising in context of the distance to uncontrollability (given in terms of the the best QMR eigenpair) is

$$\underline{C}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{25}$$

We have Jordan blocks at $\theta = 0$, $\underline{\theta} = \infty$ and $\rho = 0$.

A Jordan block: infinitely many QMR eigenvalues

A startling example used already by Eising in context of the distance to uncontrollability (given in terms of the the best QMR eigenpair) is

$$\underline{C}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{25}$$

We have Jordan blocks at $\theta = 0$, $\underline{\theta} = \infty$ and $\rho = 0$.

For $k \in \mathbb{N}$ this is an example of an infinite set of QMR eigenvalues

A Jordan block: infinitely many QMR eigenvalues

A startling example used already by Eising in context of the distance to uncontrollability (given in terms of the the best QMR eigenpair) is

$$\underline{C}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad (25)$$

We have Jordan blocks at $\theta = 0$, $\underline{\theta} = \infty$ and $\rho = 0$.

For $k \in \mathbb{N}$ this is an example of an infinite set of QMR eigenvalues,

$$\grave{\theta}_\phi = \cos\left(\frac{\pi}{k+1}\right) e^{i\phi}, \quad \phi \in [0, 2\pi). \qquad (26)$$

A Jordan block: infinitely many QMR eigenvalues

A startling example used already by Eising in context of the distance to uncontrollability (given in terms of the the best QMR eigenpair) is

$$\underline{C}_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{25}$$

We have Jordan blocks at $\theta = 0$, $\underline{\theta} = \infty$ and $\rho = 0$.

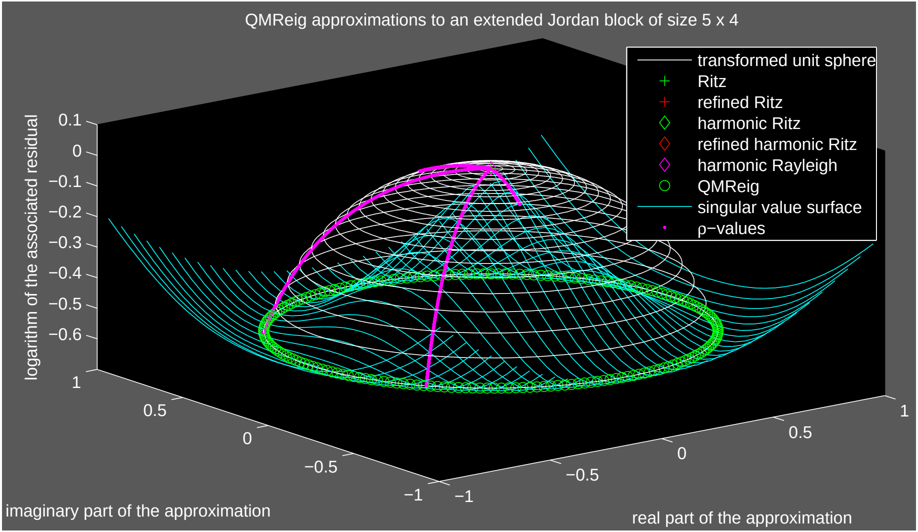
For $k \in \mathbb{N}$ this is an example of an infinite set of QMR eigenvalues,

$$\grave{\theta}_\phi = \cos\left(\frac{\pi}{k+1}\right) e^{i\phi}, \quad \phi \in [0, 2\pi). \tag{26}$$

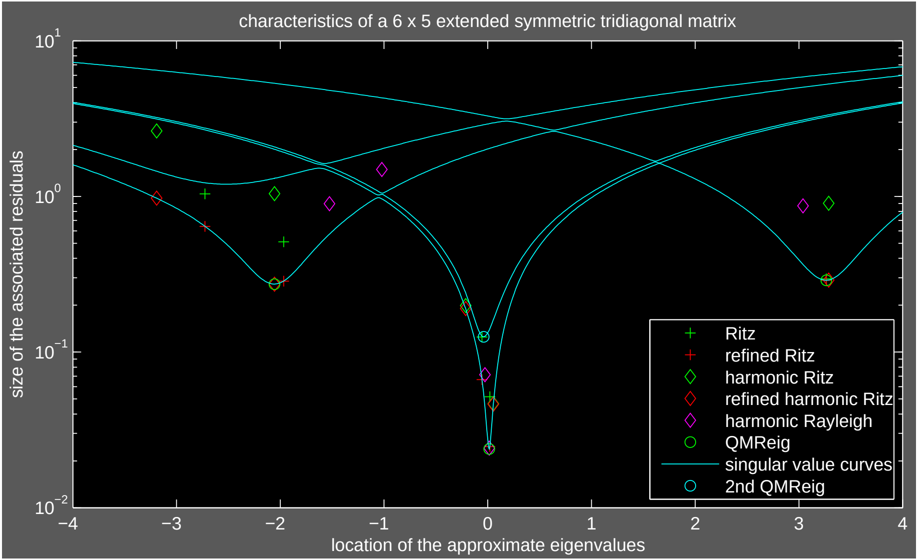
The residual of the corresponding QMR eigenpairs is given by

$$\|(\grave{\theta}_\phi \underline{I}_k - \underline{C}_k)\grave{s}_\phi\| = \sin\left(\frac{\pi}{k+1}\right) \quad \forall\ \phi. \tag{27}$$

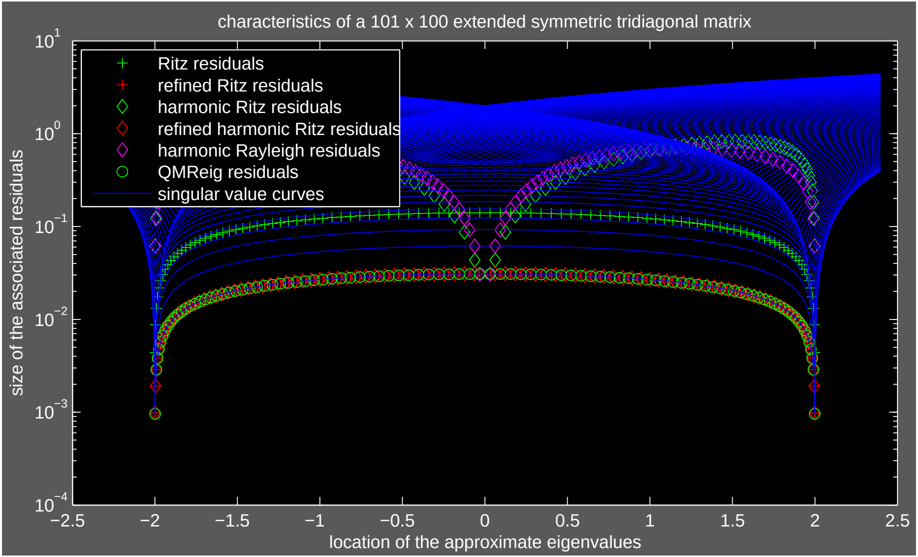
A Jordan block: infinitely many QMR eigenvalues



Always remember: It's only locally optimal . . .



Large matrices and harmonic Ritz



Lehmann's work on eigenvalues

Between 1948 and 1966 N. J. Lehmann published several papers related to "Optimale Eigenwerteinschließungen". Lehmann was interested in selfadjoint and normal linear operators (matrices).

Lehmann's work on eigenvalues

Between 1948 and 1966 N. J. Lehmann published several papers related to "Optimale Eigenwerteinschließungen". Lehmann was interested in selfadjoint and normal linear operators (matrices).

In his works we can find eigenvalue inclusions using the Temple-Quotient, shifted harmonic Ritz values, and the relation of shifted harmonic Ritz to Ritz-Galërkin, all for selfadjoint matrices.

Lehmann's work on eigenvalues

Between 1948 and 1966 N. J. Lehmann published several papers related to "Optimale Eigenwerteinschließungen". Lehmann was interested in selfadjoint and normal linear operators (matrices).

In his works we can find eigenvalue inclusions using the Temple-Quotient, shifted harmonic Ritz values, and the relation of shifted harmonic Ritz to Ritz-Galërkin, all for selfadjoint matrices.

He was interested in generalizing his results [Seite 246, 1963]:

> *"Für Aufgaben mit komplexen Eigenwerten stehen viele der Untersuchungen allerdings noch aus. Mit diesen Problemen befaßt sich eine in Arbeit befindliche Dissertation."*

Lehmann's work on eigenvalues

Between 1948 and 1966 N. J. Lehmann published several papers related to "Optimale Eigenwerteinschließungen". Lehmann was interested in selfadjoint and normal linear operators (matrices).

In his works we can find eigenvalue inclusions using the Temple-Quotient, shifted harmonic Ritz values, and the relation of shifted harmonic Ritz to Ritz-Galërkin, all for selfadjoint matrices.

He was interested in generalizing his results [Seite 246, 1963]:

> *"Für Aufgaben mit komplexen Eigenwerten stehen viele der Untersuchungen allerdings noch aus. Mit diesen Problemen befaßt sich eine in Arbeit befindliche Dissertation."*

We have extended his approach to general complex square matrices by replacing "Lehmann optimality" by "backward error". Thus, we have extended his work to normal matrices (Q_{k+1} orthonormal; Arnoldi's method).

Lehmann's results summarized

Lehmann used the information included in $Q \in \mathbb{C}^{(n,k)}$ and $W = AQ \in \mathbb{C}^{(n,k)}$, where $A \in \mathbb{C}^{(n,n)}$ is selfadjoint. We used a generic $x = Qv \in \mathbb{C}^n$.

Lehmann's results summarized

Lehmann used the information included in $Q \in \mathbb{C}^{(n,k)}$ and $W = AQ \in \mathbb{C}^{(n,k)}$, where $A \in \mathbb{C}^{(n,n)}$ is selfadjoint. We used a generic $x = Qv \in \mathbb{C}^n$.

Lehmann imposed the least-squares optimality conditions [(5a), 1963]

$$\min = \sigma^2(z) = \frac{\|(A - zI)x\|_2^2}{\|x\|_2^2} = \frac{\|(W - zQ)v\|_2^2}{\|Qv\|_2^2}$$

Lehmann's results summarized

Lehmann used the information included in $Q \in \mathbb{C}^{(n,k)}$ and $W = AQ \in \mathbb{C}^{(n,k)}$, where $A \in \mathbb{C}^{(n,n)}$ is selfadjoint. We used a generic $x = Qv \in \mathbb{C}^n$.

Lehmann imposed the least-squares optimality conditions [(5a), 1963]

$$\min = \sigma^2(z) = \frac{\|(A - zI)x\|_2^2}{\|x\|_2^2} = \frac{\|(W - zQ)v\|_2^2}{\|Qv\|_2^2}$$

and thus (by differentiation) the eigenvalue (SVD) problem [(8a), 1963]

$$Q^H(A - zI)^H(A - zI)Qv = \sigma^2(z)Q^HQv.$$

Lehmann's results summarized

Lehmann used the information included in $Q \in \mathbb{C}^{(n,k)}$ and $W = AQ \in \mathbb{C}^{(n,k)}$, where $A \in \mathbb{C}^{(n,n)}$ is selfadjoint. We used a generic $x = Qv \in \mathbb{C}^n$.

Lehmann imposed the least-squares optimality conditions [(5a), 1963]

$$\min = \sigma^2(z) = \frac{\|(A - zI)x\|_2^2}{\|x\|_2^2} = \frac{\|(W - zQ)v\|_2^2}{\|Qv\|_2^2}$$

and thus (by differentiation) the eigenvalue (SVD) problem [(8a), 1963]

$$Q^H (A - zI)^H (A - zI) Qv = \sigma^2(z) Q^H Qv.$$

Lehmann was interested in optimal shifts, i.e., shifts z resulting in a minimal radius $\sigma(z)$ of the inclusion. These are [Satz 4, 1963] among the stationary points of $\sigma^2(z)$,

$$\frac{\partial \sigma^2(z)}{\partial z} = 0.$$

Lehmann's little-known results

Differentiating an expression involving the Temple quotient $T_\tau(x)$, he obtained the shifted harmonic Ritz values [(20a)+(28), 1963] of Morgan (1991) and Freund (1992),

$$Q^H(A-\tau I)^H Qv = \frac{1}{\underline{\theta}(\tau)-\tau} Q^H(A-\tau I)^H(A-\tau I)Qv.$$

Lehmann's little-known results

Differentiating an expression involving the Temple quotient $T_\tau(x)$, he obtained the shifted harmonic Ritz values [(20a)+(28), 1963] of Morgan (1991) and Freund (1992),

$$Q^H(A-\tau I)^H Qv = \frac{1}{\underline{\theta}(\tau)-\tau} Q^H(A-\tau I)^H(A-\tau I)Qv.$$

Lehmann noticed already that poles occur in the shifted harmonic Ritz approach if using the Ritz values as shifts.

Lehmann's little-known results

Differentiating an expression involving the Temple quotient $T_\tau(x)$, he obtained the shifted harmonic Ritz values [(20a)+(28), 1963] of Morgan (1991) and Freund (1992),

$$Q^H(A-\tau I)^H Qv = \frac{1}{\underline{\theta}(\tau)-\tau} Q^H(A-\tau I)^H(A-\tau I)Qv.$$

Lehmann noticed already that poles occur in the shifted harmonic Ritz approach if using the Ritz values as shifts.

He (defined and) noted certain interesting symmetries/properties, namely

$$\tau = z \mp \sigma(z), \qquad \underline{\theta}(\tau) = z \pm \sigma(z), \qquad \text{[(Seite 251),1963]}$$

$$\underline{\theta}(\tau) = T_\tau(\underline{x}), \qquad T_\tau(\underline{x}) = \frac{\underline{x}^H(A-\tau I)^H(A-\tau I)\underline{x}}{\underline{x}^H(A-\tau I)^H\underline{x}} + \tau, \qquad \text{[(15),1963]}$$

$$2z = \tau + \underline{\theta}(\tau), \qquad z^2 - \sigma^2(z) = \tau \cdot \underline{\theta}(\tau). \qquad \text{[(8b)+(21),1963]}$$

QMR eigenpairs: Grassmann Newton

The function λ is stationary only for singular vectors for $z = v^H C_k v$. If the corresponding singular value is simple, we have found a stationary point on the corresponding singular value surface.

QMR eigenpairs: Grassmann Newton

The function λ is stationary only for singular vectors for $z = v^H C_k v$. If the corresponding singular value is simple, we have found a stationary point on the corresponding singular value surface.

The Hessean has negative eigenvalues whenever the singular value $\sigma_j(z)$ found is not a smallest one, since in forming the Hesse matrix we subtract a positive semidefinite symmetric matrix (of rank less equal two) from the realification of

$$A(z) = W^H ({}^z\underline{C}_k^H {}^z\underline{C}_k - \sigma_j(z)^2 I_k) W, \qquad W = v_j^\perp.$$

The Hermitean matrix $A(z)$ has the eigenvalues $\sigma_i^2 - \sigma_j^2$, $i \neq j$. The smallest eigenvalue of the Hesse matrix is bounded from above by Weyl's Lemma by $\lambda_{\min} \leqslant \sigma_{\min}^2 - \sigma_j^2 < 0$.

QMR eigenpairs: Grassmann Newton

The function λ is stationary only for singular vectors for $z = v^H C_k v$. If the corresponding singular value is simple, we have found a stationary point on the corresponding singular value surface.

The Hessean has negative eigenvalues whenever the singular value $\sigma_j(z)$ found is not a smallest one, since in forming the Hesse matrix we subtract a positive semidefinite symmetric matrix (of rank less equal two) from the realification of

$$A(z) = W^H({}^z\underline{C}_k^H {}^z\underline{C}_k - \sigma_j(z)^2 I_k)W, \qquad W = v_j^\perp.$$

The Hermitean matrix $A(z)$ has the eigenvalues $\sigma_i^2 - \sigma_j^2$, $i \neq j$. The smallest eigenvalue of the Hesse matrix is bounded from above by Weyl's Lemma by $\lambda_{\min} \leqslant \sigma_{\min}^2 - \sigma_j^2 < 0$.

Is every QMR eigenpair obtained through SVD minimization also obtained by Grassmannian optimization with SPD Hessean and vice versa? How to prove this?

QMR eigenpairs: Grassmann Newton

One direction is quite simple: As any QMR eigenpair $(\grave{\theta}, \grave{v})$ from the SVD minimization satisfies

$$\grave{\theta} = \frac{\grave{v}^H C_k \grave{v}}{\grave{v}^H \grave{v}} = z(\grave{v}),$$

the minimum is obtained as a function value of the function λ living on the first Grassmannian.

QMR eigenpairs: Grassmann Newton

One direction is quite simple: As any QMR eigenpair $(\grave{\theta}, \grave{v})$ from the SVD minimization satisfies

$$\grave{\theta} = \frac{\grave{v}^H C_k \grave{v}}{\grave{v}^H \grave{v}} = z(\grave{v}),$$

the minimum is obtained as a function value of the function λ living on the first Grassmannian.

There can be no smaller function value nearby, since this would result in a sequence of $v(\epsilon_i)$, $\|v(\epsilon_i)\| = 1$, arbitrarily close to $v = \grave{v}$, $\|v\| = 1$, with

$$\sigma_k(z(v)) = \|(z(v)\underline{I}_k - \underline{C}_k)v\| > \|(z(v(\epsilon_i))\underline{I}_k - \underline{C}_k)v(\epsilon_i)\| \geqslant \sigma_k(z(v(\epsilon_i))),$$

thus, there would be a sequence $z_i = z(v(\epsilon_i)) \to z = \grave{\theta}$ with

$$\sigma_k(\grave{\theta}) > \sigma_k(z_i),$$

which gives with the continuity of the functions involved a contradiction.

QMR eigenpairs: Grassmann Newton

If the singular vector to the smallest singular values as a function of the parameter ϵ is continuous, which is at least the case if the singular value is simple and thus real analytic by Sun's results, we can prove that a minimum of the vector-valued function is indeed also a minimum of the singular value surface.

QMR eigenpairs: Grassmann Newton

If the singular vector to the smallest singular values as a function of the parameter ϵ is continuous, which is at least the case if the singular value is simple and thus real analytic by Sun's results, we can prove that a minimum of the vector-valued function is indeed also a minimum of the singular value surface.

Denote the "smallest" singular vector of ${}^{(z+\epsilon)}\underline{C}_k$ by $v(\epsilon) = v_k(z+\epsilon)$, i.e.,

$$((z+\epsilon)\underline{I}_k - \underline{C}_k)v(\epsilon) = u_k(z+\epsilon)\sigma_k(z+\epsilon),$$

and define

$$z(\epsilon) = \frac{v(\epsilon)^H C_k v(\epsilon)}{v(\epsilon)^H v(\epsilon)}.$$

QMR eigenpairs: Grassmann Newton

If the singular vector to the smallest singular values as a function of the parameter ϵ is continuous, which is at least the case if the singular value is simple and thus real analytic by Sun's results, we can prove that a minimum of the vector-valued function is indeed also a minimum of the singular value surface.

Denote the "smallest" singular vector of ${}^{(z+\epsilon)}\underline{C}_k$ by $v(\epsilon) = v_k(z+\epsilon)$, i.e.,

$$((z+\epsilon)\underline{I}_k - \underline{C}_k)v(\epsilon) = u_k(z+\epsilon)\sigma_k(z+\epsilon),$$

and define

$$z(\epsilon) = \frac{v(\epsilon)^H C_k v(\epsilon)}{v(\epsilon)^H v(\epsilon)}.$$

Then, obviously,

$$\sigma_k^2(z+\epsilon) \geqslant \|{}^{z(\epsilon)}\underline{C}_k v(\epsilon)\|_2^2 \geqslant \sigma_k^2(z(\epsilon)). \tag{28}$$

QMR eigenpairs: Grassmann Newton

We now use the "closeness" of the functions σ_k^2 and λ, i.e.,

$$\sigma_k^2(z+\epsilon) \geqslant \lambda(v(\epsilon)) = \|{}^{z(\epsilon)}\underline{C}_k v(\epsilon)\|_2^2 \geqslant \sigma_k^2(z(\epsilon)). \tag{29}$$

QMR eigenpairs: Grassmann Newton

We now use the “closeness” of the functions σ_k^2 and λ, i.e.,

$$\sigma_k^2(z+\epsilon) \geqslant \lambda(v(\epsilon)) = \|{}^{z(\epsilon)}\underline{C}_k v(\epsilon)\|_2^2 \geqslant \sigma_k^2(z(\epsilon)). \tag{29}$$

Suppose that we do not have a minimum at z, i.e., let N_z denote a neighborhood of z,

$$\forall N_z \quad \exists w \in N_z \, : \, \sigma_k(w) < \sigma_k(z), \tag{30}$$

or, by the Axiom of Choice, a sequence of points $\{\epsilon_i \in \mathbb{C}\}_{i=1}^{\infty}$ such that

$$\lim_{i\to\infty} \epsilon_i = 0 \quad \text{and} \quad \sigma_k(z+\epsilon_i) < \sigma_k(z) \quad \forall i \in \mathbb{N}. \tag{31}$$

QMR eigenpairs: Grassmann Newton

We now use the “closeness” of the functions σ_k^2 and λ, i.e.,

$$\sigma_k^2(z+\epsilon) \geqslant \lambda(v(\epsilon)) = \|{}^{z(\epsilon)}\underline{C}_k v(\epsilon)\|_2^2 \geqslant \sigma_k^2(z(\epsilon)). \tag{29}$$

Suppose that we do not have a minimum at z, i.e., let N_z denote a neighborhood of z,

$$\forall N_z \quad \exists w \in N_z \,:\, \sigma_k(w) < \sigma_k(z), \tag{30}$$

or, by the Axiom of Choice, a sequence of points $\{\epsilon_i \in \mathbb{C}\}_{i=1}^{\infty}$ such that

$$\lim_{i\to\infty} \epsilon_i = 0 \quad \text{and} \quad \sigma_k(z+\epsilon_i) < \sigma_k(z) \quad \forall i \in \mathbb{N}. \tag{31}$$

Choosing $i > i_0$ such that $v(\epsilon)$ is continuous we know by (29) and (31) that

$$\|{}^{z}\underline{C}_k v\|_2^2 = \sigma_k^2(z) > \sigma_k^2(z+\epsilon_i) \geqslant \|{}^{z(\epsilon_i)}\underline{C}_k v(\epsilon_i)\|_2^2 \quad \forall i > i_0. \tag{32}$$

QMR eigenpairs: Grassmann Newton

We now use the “closeness” of the functions σ_k^2 and λ, i.e.,

$$\sigma_k^2(z+\epsilon) \geqslant \lambda(v(\epsilon)) = \|{}^{z(\epsilon)}\underline{C}_k v(\epsilon)\|_2^2 \geqslant \sigma_k^2(z(\epsilon)). \tag{29}$$

Suppose that we do not have a minimum at z, i.e., let N_z denote a neighborhood of z,

$$\forall N_z \quad \exists w \in N_z \,:\, \sigma_k(w) < \sigma_k(z), \tag{30}$$

or, by the Axiom of Choice, a sequence of points $\{\epsilon_i \in \mathbb{C}\}_{i=1}^{\infty}$ such that

$$\lim_{i\to\infty} \epsilon_i = 0 \quad \text{and} \quad \sigma_k(z+\epsilon_i) < \sigma_k(z) \quad \forall i \in \mathbb{N}. \tag{31}$$

Choosing $i > i_0$ such that $v(\epsilon)$ is continuous we know by (29) and (31) that

$$\|{}^{z}\underline{C}_k v\|_2^2 = \sigma_k^2(z) > \sigma_k^2(z+\epsilon_i) \geqslant \|{}^{z(\epsilon_i)}\underline{C}_k v(\epsilon_i)\|_2^2 \quad \forall i > i_0. \tag{32}$$

This gives a contradiction.

Harmonic Ritz and ρ-values

It can be shown that the shifted harmonic Ritz values $\underline{\theta}$, the shift ("target") τ and the resulting ρ-values are related.

Harmonic Ritz and ρ-values

It can be shown that the shifted harmonic Ritz values $\underline{\theta}$, the shift ("target") τ and the resulting ρ-values are related.

In non-homogeneous description with zero shift this relation is given by

$$\|\underline{C}_k \underline{v}_j\|_2^2 = \underline{\theta}_j \overline{\rho_j} \in \mathbb{R}_+, \qquad (33)$$

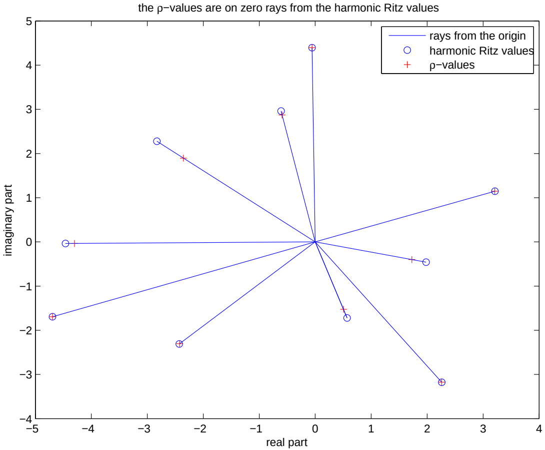
Harmonic Ritz and ρ-values

It can be shown that the shifted harmonic Ritz values $\underline{\theta}$, the shift (“target”) τ and the resulting ρ-values are related.

In non-homogeneous description with zero shift this relation is given by

$$\|\underline{C}_k \underline{v}_j\|_2^2 = \underline{\theta}_j \overline{\rho_j} \in \mathbb{R}_+, \qquad (33)$$

i.e., both $\underline{\theta}_j$ and ρ_j are on the same ray originating from zero and on the same side.



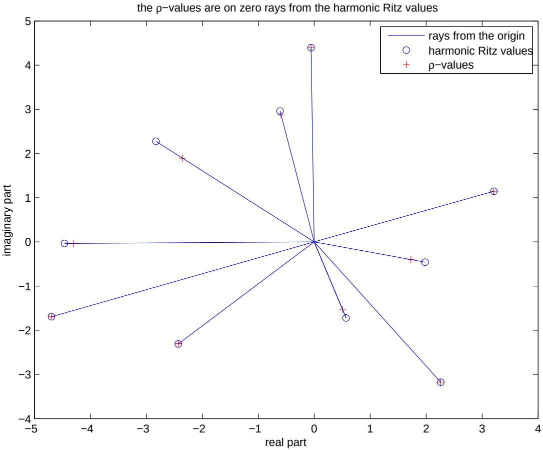
Harmonic Ritz and ρ-values

It can be shown that the shifted harmonic Ritz values $\underline{\theta}$, the shift ("target") τ and the resulting ρ-values are related.

In non-homogeneous description with zero shift this relation is given by

$$\|\underline{C}_k \underline{v}_j\|_2^2 = \underline{\theta}_j \overline{\rho_j} \in \mathbb{R}_+, \qquad (33)$$

i.e., both $\underline{\theta}_j$ and ρ_j are on the same ray originating from zero and on the same side.



In the general setting the ρ-values and the harmonic Ritz values $\underline{\theta}$ are, again, on the same rays, but now originating from the target $\tau \in \mathbb{C}$.