

Majorization-based convergence rate bounds for subspace iterations and the block Lanczos method

Merico Argentati (speaker) and Andrew Knyazev

Department of Mathematical & Statistical Sciences

University of Colorado Denver

Supported by the National Science Foundation

10th Copper Mountain Conference on Iterative Methods, 2008

April 9, 2008

Outline

- The Rayleigh–Ritz method, principal angles between subspaces \mathcal{X} and \mathcal{Y} , majorization
- Weak majorization bounds on the change in Ritz values
- Bound on principal angles between \mathcal{X} and $F\mathcal{Y}$ in terms of the principal angles between \mathcal{X} and \mathcal{Y} where F is an operator
- Subspace iteration and weak majorization bounds on Ritz values
- Convergence rate bounds for the block Lanczos method
- Conclusions

The Rayleigh–Ritz Method

- Let A be a Hermitian matrix and \mathcal{X} a subspace
- We define an operator $P_{\mathcal{X}}A|_{\mathcal{X}}$ on \mathcal{X} , where $P_{\mathcal{X}}$ is the orthogonal projection onto \mathcal{X} and $P_{\mathcal{X}}A|_{\mathcal{X}}$ denotes the restriction of $P_{\mathcal{X}}A$ to \mathcal{X} , as discussed in Parlett [1998]. The eigenvalues $\Lambda(P_{\mathcal{X}}A|_{\mathcal{X}})$ are called Ritz values.
- We have $\Lambda(P_{\mathcal{X}}A|_{\mathcal{X}}) = \Lambda(X^HAX)$ where X is a matrix with orthonormal columns that span \mathcal{X}
- Nonzero Ritz values are the nonzero eigenvalues of $P_{\mathcal{X}}AP_{\mathcal{X}}$

Bounding Eigenvalues

If A is a Hermitian matrix, we have Weyl's theorem

$$\max_{j=1,\dots,n} |\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|, \quad (1)$$

and by an analog of the Hoffman–Wielandt theorem, e.g. Stewart and Sun [1990], we have

$$\sqrt{\sum_{j=1}^n (\lambda_j(A) - \lambda_j(B))^2} \leq \|A - B\|_F \quad (2)$$

How do we bound Ritz values when we vary/perturb the subspace?

Changes in the Trial Subspace in the Rayleigh–Ritz Method

- We can vary a Hermitian matrix and see how the eigenvalues change
- Analogously for a fixed Hermitian matrix, we can vary the subspace and see how the Ritz values change. If a subspace \mathcal{X} is perturbed to give rise to another subspace \mathcal{Y} , then we can ask about how to bound

$$|\Lambda(P_{\mathcal{X}}A|_{\mathcal{X}}) - \Lambda(P_{\mathcal{Y}}A|_{\mathcal{Y}})| ?$$

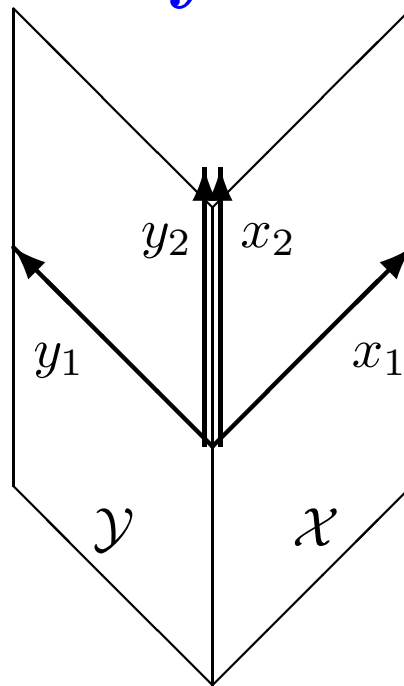
- **Answer:** We can prove some very flexible and useful bounds using principal angles between subspaces and majorization, also considering the case when \mathcal{X} or \mathcal{Y} may be invariant relative to A .

Principal Angles Between Subspaces

- Let subspaces \mathcal{X} and $\mathcal{Y} \subseteq \mathbf{C}^n$ with $\dim \mathcal{X} = \dim \mathcal{Y}$, and orthonormal bases given by the columns of the matrices X and Y . The principal angles between \mathcal{X} and \mathcal{Y} arranged in descending order are given by $\Theta(\mathcal{X}, \mathcal{Y}) = \Theta^\downarrow(\mathcal{X}, \mathcal{Y})$, and defined using $\cos \Theta(\mathcal{X}, \mathcal{Y}) = S^\uparrow(X^H Y)$
- Definition is symmetric: $\Theta(\mathcal{X}, \mathcal{Y}) = \Theta(\mathcal{Y}, \mathcal{X})$.
- Definition of the distance between subspaces:
 $\text{gap}(\mathcal{X}, \mathcal{Y}) = \|P_{\mathcal{X}} - P_{\mathcal{Y}}\|_2 = \sin(\Theta_{\max}(\mathcal{X}, \mathcal{Y}))$.

Pioneering results using angles between subspaces in the framework of unitarily invariant norms and symmetric gauge functions, equivalent to majorization, appear in Davis and Kahan [1970], which introduces many of the tools that we use here.

The Geometry of Principal Angles



$$\theta_1 \geq \dots \geq \theta_m$$

The cosines can be defined recursively for $k = m, \dots, 1$ by
 $\cos(\theta_k) = \max_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} x^T y = x_k^T y_k$ subject to
 $\|x\| = \|y\| = 1, \quad x^T x_i = 0, \quad y^T y_i = 0, \quad i = m, \dots, k + 1.$

Majorization and Basic Notation

- For vector $x = [x_1, \dots, x_n]$, we use $x^\downarrow \equiv [x_1^\downarrow, \dots, x_n^\downarrow]$ to denote x with its elements rearranged in descending order, while $x^\uparrow \equiv [x_1^\uparrow, \dots, x_n^\uparrow]$ denotes x with its elements rearranged in ascending order. $|x|$ denotes the vector x with the absolute value of its components.
- We say that $x \in \mathcal{R}^n$ is **weakly majorized** by $y \in \mathcal{R}^n$, written $x \prec_w y$, if

$$\sum_{i=1}^k x_i^\downarrow \leq \sum_{i=1}^k y_i^\downarrow, \quad 1 \leq k \leq n, \quad (3)$$

while x is (strongly) **majorized** by y , written $x \prec y$, if (3) holds together with

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i. \quad (4)$$

Why is Majorization Important?

- The notion of majorization plays an important role in mathematics, statistics and economics
- The following two conditions are equivalent: $x \prec_w y$ and $\sum_{i=1}^n \phi(x_i) \leq \sum_{i=1}^n \phi(y_i)$ for all nondecreasing convex functions ϕ . This implies inequalities for unitarily invariant norms and certain p -norms
- For Hermitian matrices A and B let $\Lambda(A)$ and $S(A)$ be, respectively, the vector of eigenvalues and singular values (both nonincreasing), then

$$\text{diag}(A) \prec \Lambda(A), \quad (\text{Schur's Theorem}) \quad (5)$$

$$\Lambda(A) - \Lambda(B) \prec \Lambda(A - B) \quad \text{and} \quad |\Lambda(A) - \Lambda(B)| \prec_w S(A - B) \quad (6)$$

See Lidskii [1950], Mirsky [1960], Marshall and Olkin [1979], Horn and Johnson [1999].

Outline

- The Rayleigh–Ritz method, principal angles between subspaces \mathcal{X} and \mathcal{Y} , majorization
- Weak majorization bounds on the change in Ritz values
- Bound on principal angles between \mathcal{X} and $F\mathcal{Y}$ in terms of the principal angles between \mathcal{X} and \mathcal{Y} where F is an operator
- Subspace iteration and weak majorization bounds on Ritz values
- Convergence rate bounds for the block Lanczos method
- Conclusions

Changes in the Trial Subspace in the Rayleigh–Ritz Method

The result of (Argentati [2003], Knyazev and Argentati [2006a]):

Theorem 1 *Let A be a Hermitian matrix and let \mathcal{X} and \mathcal{Y} be subspaces with $\dim\mathcal{X} = \dim\mathcal{Y} = m$. Then*

$$\max_{j=1,\dots,m} |\lambda_j(P_{\mathcal{X}}A|_{\mathcal{X}}) - \lambda_j(P_{\mathcal{Y}}A|_{\mathcal{Y}})| \leq (\lambda_{\max} - \lambda_{\min}) \sin(\Theta_{\max}(\mathcal{X}, \mathcal{Y}))$$

where λ_{\min} and λ_{\max} are the smallest and largest eigenvalues of A .

Main Weak Majorization Theorem for the Rayleigh–Ritz Method

Theorem 2 *Let A be a Hermitian matrix and let \mathcal{X} and \mathcal{Y} be subspaces with $\dim\mathcal{X} = \dim\mathcal{Y}$, then (see Knyazev and Argentati [2006b])*

$$|\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) - \Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}})| \prec_w (\lambda_{\max} - \lambda_{\min}) \sin \Theta(\mathcal{X}, \mathcal{Y}), \quad (7)$$

where λ_{\min} and λ_{\max} are the smallest and largest eigenvalues of the A , respectively. If in addition one of the subspaces is A -invariant then (see conjecture in Argentati, Knyazev, Paige, and Panayotov [2008])

$$|\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) - \Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}})| \prec_w (\lambda_{\max} - \lambda_{\min}) \sin^2 \Theta(\mathcal{X}, \mathcal{Y}). \quad (8)$$

Theorem 2 is Sharp

Consider the following example for (8) where \mathcal{X} is A -invariant (see Argentati, Knyazev, Paige, and Panayotov [2008]). Let I be an $m \times m$ identity matrix and let $A = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$. Let Σ be a diagonal of given

cosines and let Γ be the diagonal of sines, so $\Sigma^2 + \Gamma^2 = I$. Let $X = \begin{bmatrix} I \\ 0 \end{bmatrix}$,

and $Y = \begin{bmatrix} \Sigma \\ \Gamma \end{bmatrix}$. Then $X^H AX = I$ and $Y^H AY = 2\Sigma^2 - I$ and

$$|\Lambda(X^H AX) - \Lambda(Y^H AY)|^\downarrow = 2[\sin^2 \Theta(\mathcal{X}, \mathcal{Y})] = (\lambda_{\max} - \lambda_{\min})[\sin^2 \Theta(\mathcal{X}, \mathcal{Y})].$$

Weak Majorization Implications of

$$|\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) - \Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}})| \prec_w (\lambda_{\max} - \lambda_{\min}) \sin^2 \Theta(\mathcal{X}, \mathcal{Y})$$

The implications of weak majorization may not be obvious. Let $\dim \mathcal{X} = \dim \mathcal{Y} = m$, and let $\alpha_1 \geq \dots \geq \alpha_m$ be the Ritz values of A with respect to \mathcal{X} and $\beta_1 \geq \dots \geq \beta_m$ be the Ritz values of A with respect to \mathcal{Y} . Then using (8) we have

$$\sum_{i=1}^k |\alpha_i - \beta_i|^{\downarrow} \leq (\lambda_{\max} - \lambda_{\min}) \sum_{i=1}^k \sin^2(\Theta_i(\mathcal{X}, \mathcal{Y})), \quad k = 1, \dots, m.$$

For $k = 1$ we have

$$\begin{aligned} \max_{j=1, \dots, m} |\alpha_j - \beta_j| &\leq (\lambda_{\max} - \lambda_{\min}) \sin^2(\Theta_{\max}(\mathcal{X}, \mathcal{Y})) \\ &= (\lambda_{\max} - \lambda_{\min}) \text{gap}^2(\mathcal{X}, \mathcal{Y}), \end{aligned}$$

Weak Majorization Implications (Cont.)

$$|\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) - \Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}})| \prec_w (\lambda_{\max} - \lambda_{\min}) \sin^2 \Theta(\mathcal{X}, \mathcal{Y}),$$

and for $k = m$ we obtain

$$\sum_{i=1}^m |\alpha_i - \beta_i| \leq (\lambda_{\max} - \lambda_{\min}) \sum_{i=1}^m \sin^2(\Theta_i(\mathcal{X}, \mathcal{Y}))$$

We have in general

$$\left(\sum_{i=1}^m |\alpha_i - \beta_i|^p \right)^{\frac{1}{p}} \leq (\lambda_{\max} - \lambda_{\min}) \left(\sum_{i=1}^m \sin^{2p}(\Theta_i(\mathcal{X}, \mathcal{Y})) \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty.$$

Bounding Eigenvalues (Cont.)

If A is a Hermitian matrix, we have Weyl's theorem

$$\max_{j=1,\dots,n} |\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|, \quad (9)$$

and by an analog of the Hoffman–Wielandt theorem, e.g. Stewart and Sun [1990], we have

$$\sqrt{\sum_{j=1}^n (\lambda_j(A) - \lambda_j(B))^2} \leq \|A - B\|_F \quad (10)$$

Note that the above formulas are two of a large class of inequalities that follow from

$$|\Lambda(A) - \Lambda(B)| \prec_w S(A - B)$$

Matlab Example for Many Zero Angles

Let $n = 5$ and $\dim \mathcal{X} = \dim \mathcal{Y} = 4$

$$A = \text{diag}([1 \ .9 \ .7 \ .5 \ 0]) \quad X = \text{eye}(n, m) \quad Y = \text{orth}(\text{rand}(n, m)).$$

Then \mathcal{X} is A -invariant and

$$\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) = [1.0000 \ 0.9000 \ 0.7000 \ 0.5000],$$

$$\Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}}) = [0.9452 \ 0.7484 \ 0.6572 \ 0.0330].$$

Then

$$|\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) - \Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}})|^{\downarrow} = [0.4670 \ 0.1516 \ 0.0548 \ 0.0428]$$

$$\prec_w (\lambda_{\max} - \lambda_{\min}) \sin^2 \Theta(\mathcal{X}, \mathcal{Y}) = [0.9513 \ 0 \ 0 \ 0].$$

Conclusion: Zero angles do not imply that $\Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}})$ includes eigenvalues of A .

Outline

- The Rayleigh–Ritz method, principal angles between subspaces \mathcal{X} and \mathcal{Y} , majorization
- Weak majorization bounds on the change in Ritz values
- Bound on principal angles between \mathcal{X} and $F\mathcal{Y}$ in terms of the principal angles between \mathcal{X} and \mathcal{Y} where F is an operator
- Subspace iteration and weak majorization bounds on Ritz values
- Convergence rate bounds for the block Lanczos method
- Conclusions

Subspace Iterations: $F\mathcal{Y}$ as a Possibly Improved Approximation of \mathcal{X}

Theorem 3 *Let $\dim\mathcal{X} = \dim\mathcal{Y}$ and assume $\Theta(\mathcal{X}, \mathcal{Y}) < \pi/2$. Let F be invariant on both \mathcal{X} and on its orthogonal complement \mathcal{X}^\perp , and assume that $(P_{\mathcal{X}}F)|_{\mathcal{X}}$ is invertible. Then $\dim(F\mathcal{Y}) = \dim\mathcal{Y}$ and*

$$\tan \Theta(\mathcal{X}, F\mathcal{Y}) \prec_w S \left(((P_{\mathcal{X}}F)|_{\mathcal{X}})^{-1} \right) S \left((P_{\mathcal{X}^\perp}F)|_{\mathcal{X}^\perp} \right) \tan \Theta(\mathcal{X}, \mathcal{Y}).$$

We can also prove the stronger form

$$\log \left(\frac{\tan \Theta(\mathcal{X}, F\mathcal{Y})}{\tan \Theta(\mathcal{X}, \mathcal{Y})} \right) \prec_w \log \left(S \left(((P_{\mathcal{X}}F)|_{\mathcal{X}})^{-1} \right) S \left((P_{\mathcal{X}^\perp}F)|_{\mathcal{X}^\perp} \right) \right).$$

Subspace Iterations and Ritz Values

Theorem 4 *Let \mathcal{X} and \mathcal{Y} be subspaces of \mathcal{H} such that $\dim\mathcal{X} = \dim\mathcal{Y}$ and $\Theta(\mathcal{X}, \mathcal{Y}) < \pi/2$. Let the operator A be Hermitian, and let \mathcal{X} be an A -invariant subspace. Let $F = f(A)$ and $f(\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}})) \neq 0$. Then $\dim(F\mathcal{Y}) = \dim\mathcal{Y}$ and*

$$|\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) - \Lambda((P_{F\mathcal{Y}}A)|_{F\mathcal{Y}})| \\ \prec_w (\lambda_{\max} - \lambda_{\min}) \left(\frac{|f(\Lambda((P_{\mathcal{X}^\perp}A)|_{\mathcal{X}^\perp}))|^\downarrow}{|f(\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}))|^\uparrow} \right)^2 \tan^2 \Theta(\mathcal{X}, \mathcal{Y}).$$

Subspace Iterations and Ritz Values

Example: $f(A) = A^k$

Let the A -invariant subspace X correspond to the contiguous set of the largest eigenvalues of A . Let

$$\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) = [\lambda_1 \geq \dots \geq \lambda_m] \quad \text{top } m \text{ eigenvalues of } A$$

$$\Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}}) = [\tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_m].$$

Then $\lambda_i \geq \tilde{\lambda}_i$, $i = 1, \dots, m$ and

$$\begin{aligned} & [\lambda_1 - \tilde{\lambda}_1, \dots, \lambda_m - \tilde{\lambda}_m] \\ \prec_w (\lambda_{\max} - \lambda_{\min}) & \left[\left(\frac{\lambda_{m+1}}{\lambda_m} \right)^{2k} \tan^2 \theta_1(\mathcal{X}, \mathcal{Y}), \dots, \left(\frac{\lambda_{2m}}{\lambda_1} \right)^{2k} \tan^2 \theta_m(\mathcal{X}, \mathcal{Y}) \right] \end{aligned}$$

Outline

- The Rayleigh–Ritz method, principal angles between subspaces \mathcal{X} and \mathcal{Y} , majorization
- Weak majorization bounds on the change in Ritz values
- Bound on principal angles between \mathcal{X} and $F\mathcal{Y}$ in terms of the principal angles between \mathcal{X} and \mathcal{Y} where F is an operator
- Subspace iteration and weak majorization bounds on Ritz values
- **Convergence rate bounds for the block Lanczos method**
- Conclusions

Convergence Rate Bounds of the Block Lanczos Method

Theorem 5 *Let $\dim \mathcal{X} = \dim \mathcal{X}_0 = m$, operator A be Hermitian, and let the A -invariant subspace X correspond to the contiguous set of the largest eigenvalues of A . Let $\mathcal{Y} = \mathcal{X}_0 + A\mathcal{X}_0 + \dots + A^k \mathcal{X}_0$, then*

$$0 \leq \frac{\Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) - \Lambda_{\dim \mathcal{X}}((P_{\mathcal{Y}}A)|_{\mathcal{Y}})}{\Lambda_{\dim \mathcal{X}}((P_{\mathcal{Y}}A)|_{\mathcal{Y}}) - \lambda_{\min}} \prec_w [\sigma_m^*, \dots, \sigma_1^*] \tan^2 \Theta(\mathcal{X}, \mathcal{X}_0).$$

where

$$\sigma_i^* = \left| T_k \left(\frac{\lambda_{m+1} + \lambda_{\min} - 2\lambda_i}{\lambda_{m+1} - \lambda_{\min}} \right) \right|^{-2} \quad i = 1, \dots, m,$$

and where T_k is the k th Chebyshev polynomial of the first kind.

See, e.g., Saad [1980], Knyazev [1987].

Convergence Rate Bounds of the Block Lanczos Method (Cont.)

Here is a more concrete interpretation of Theorem 5. Let

$$\begin{aligned} \Lambda((P_{\mathcal{X}}A)|_{\mathcal{X}}) &= [\lambda_1 \geq \dots \geq \lambda_m] \quad \text{top } m \text{ eigenvalues of } A \\ \Lambda_{\dim \mathcal{X}}((P_{\mathcal{Y}}A)|_{\mathcal{Y}}) &= [\tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_m] \quad \text{top } m \text{ Ritz values of } \Lambda((P_{\mathcal{Y}}A)|_{\mathcal{Y}}) \end{aligned}$$

Then

$$\begin{aligned} &\left[\left(\frac{\lambda_1 - \tilde{\lambda}_1}{\tilde{\lambda}_1 - \lambda_{\min}} \right), \dots, \left(\frac{\lambda_m - \tilde{\lambda}_m}{\tilde{\lambda}_m - \lambda_{\min}} \right) \right] \\ &\prec_w [\sigma_m^* \tan^2 \theta_1(\mathcal{X}, \mathcal{X}_0), \dots, \sigma_1^* \tan^2 \theta_m(\mathcal{X}, \mathcal{X}_0)] \end{aligned}$$

Conclusions

- We can prove some very flexible and useful bounds for Ritz values using principal angles and majorization
- The absolute value of the difference of Ritz values for a Hermitian matrix are majorized by $\lambda_{\max} - \lambda_{\min}$ times the sines of the angles between the perturbed subspaces, with majorization by the squares of the sines (see conjecture in [2008]) if one of the subspaces is A -invariant
- We can characterize a possible improved approximation to X by applying a function F to Y , in terms of singular values and the tangent of principal angles
- Using this general approach we obtain a useful characterization of the eigenvalue error for the Block Lanczos Method

References

- M. E. Argentati, A. V. Knyazev, C. C. Paige, and I. Panayotov. Bounds on changes in Ritz values for a perturbed invariant subspace of a Hermitian matrix. *SIAM J. Matrix Anal. Appl.* Accepted, 2008.
- Merico E. Argentati. *Principal Angles Between Subspaces as Related to Rayleigh Quotient and Rayleigh Ritz Inequalities with Applications to Eigenvalue Accuracy and an Eigenvalue Solver*. PhD thesis, University of Colorado at Denver, 2003.
- C. Davis and W. M. Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.
- R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, New York, NY, 1999.
- A. V. Knyazev. Convergence rate estimates for iterative methods for mesh symmetric eigenvalue problem. *Soviet J. Numerical Analysis and Math. Modelling*, 2(5):371–396, 1987.
- A. V. Knyazev and M. E. Argentati. On proximity of Rayleigh quotients for different vectors and Ritz values generated by different trial subspaces. *Linear Algebra and its Applications*, 415:82–95, 2006a.

- A. V. Knyazev and M. E. Argentati. Majorization for changes in angles between subspaces, Ritz values, and graph Laplacian spectra. *SIAM J. Matr. Anal. Appl.*, 29:15–32, 2006b.
- V. B. Lidskii. On the proper values of a sum and product of symmetric matrices. *Dokl. Akad. Nauk SSSR*, 75:769–772, 1950.
- A. W. Marshall and I. Olkin. *Inequalities: theory of majorization and its applications*, volume 143 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979. ISBN 0-12-473750-1.
- L. Mirsky. Symmetric gauge functions and unitarily invariant norms. *Quarterly Journal of Mathematics. Oxford Series(2)*, 11:50–59, 1960.
- Beresford N. Parlett. *The Symmetric Eigenvalue Problem*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. ISBN 0-89871-402-8. Corrected reprint of the 1980 original.
- Y. Saad. On the rates of convergence of the lanczos and the block–lanczos methods. *SIAM Journal on Numerical Analysis*, 17(5):687–706, 1980.
- G. W. Stewart and J. G. Sun. *Matrix perturbation theory*. Academic Press Inc., Boston, MA, 1990. ISBN 0-12-670230-6.

Backup Slides

We Can Replace $\lambda_{\max} - \lambda_{\min}$ With a Reduced Constant

- As in Knyazev and Argentati [2006a, Remark 7], the constant $\lambda_{\max} - \lambda_{\min}$ can be replaced with

$$\max_{x \in \mathcal{X} + \mathcal{Y}, \|x\|=1} (x, Ax) - \min_{x \in \mathcal{X} + \mathcal{Y}, \|x\|=1} (x, Ax),$$

which for some subspaces \mathcal{X} and \mathcal{Y} can provide a significant improvement.

- This effectively replaces A with $P_{(\mathcal{X} + \mathcal{Y})} A P_{(\mathcal{X} + \mathcal{Y})}$, which has the same action on \mathcal{X} and \mathcal{Y} .