Ritz Value Bounds That Exploit Quasi-Sparsity

Ilse Ipsen

North Carolina State University

Motivation

- Quantum Physics
- Small eigenvalues of large Hamiltonians
- Quasi-Sparse Eigenvector Method (QSE) [Lee, Salwen & Lee 2001]
- QSE computes Ritz values of leading principal submatrices
- Ritz values surprisingly accurate
- Why?

Example

$$H = \begin{pmatrix} 100 & \epsilon \\ \epsilon & 1 \end{pmatrix} \qquad 0 \le \epsilon < 1$$

QSE approximates $\lambda_1 \approx 1$ by Ritz value 100

We want: $|100 - \lambda_1| \le ?$

But Ritz value bound gives only:

$$|100 - \lambda_2| \le \epsilon \qquad (\lambda_2 \approx 100)$$

Existing Ritz Value Bounds

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} \quad \text{Hermitian}$$

There exist eigenvalues λ_{k_j} of H such that

$$|\lambda_j(H_{11}) - \lambda_{k_j}| \leq \|H_{12}\|$$

- No control over which eigenvalues approximated
- No exploitation of small eigenvector elements

Literature: Hermitian Matrices

- Kahan 1967, Parlett 1980
  Residual bounds for clustered Ritz values
- Kaniel 1966, Saad 1980,
  Sleijpen, Van Den Eshof & Smit 2002
  A priori bounds in terms of subspace angles
- Drmač & Hari, 1997
  Relative bounds for semi-definite matrices
- Kuijlaars, 2000
  Asymptotic bounds, potential theory

Literature: Non-Hermitian Matrices

- Jia & Stewart 2001
  A priori bounds in terms of subspace angles
- A priori bounds for Krylov space methods (Lanczos, GMRES, FOM)

Disadvantages

- Bounds don't target specific eigenvalues
- Bounds don't exploit structure of eigenvectors

Overview

- (Algebraically) Smallest Eigenvalue
  - Quasi-Sparsity
  - Hermitian Matrices
  - Hermitian Tridiagonal Matrices
  - Relative Bound for Hermitian Matrices
- Several Eigenvalues
  - (Partially) Hermitian Matrices
  - Real Eigenvalues or Ritz Values
  - General Complex Matrices

Quasi-Sparsity

$$Hv = \lambda v$$

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Quasi-sparsity of eigenvector v

$$\rho \equiv \frac{||v_2||}{||v_1||} \qquad \text{(two norm)}$$

If $v_1 = 0$ then $\rho = \infty$

Hermitian Ritz Value Bounds

$Hv = \lambda_1 v$ $\quad \lambda_1$ smallest eigenvalue, distinct

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

For smallest eigenvalue θ_1 of H_{11}:

$$0 \leq \theta_1 - \lambda_1 \leq \|H_{12}\|\,\rho \qquad \rho \equiv \frac{\|v_2\|}{\|v_1\|}$$

Example

$$H = \begin{pmatrix} 100 & \epsilon \\ \epsilon & 1 \end{pmatrix} \qquad \rho \leq \frac{99+\epsilon}{\epsilon}$$

$$0 \leq 100 - \lambda_1 \leq 99 + \epsilon$$

Upper bound as good as Weyl's theorem:

$$99 - \epsilon \leq 100 - \lambda_1 \leq 99 + \epsilon$$

Hermitian Matrices

$Hv = \lambda_1 v$

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

If $\|v_2\| \ll \|v_1\|$ then

$$|\theta_1 - \lambda_1| \ll \|H_{12}\|$$

Smallest Ritz value good approximation to smallest eigenvalue, if eigenvector quasi-sparse

Hermitian Tridiagonals

$$Tv = \lambda_1 v$$

$$T = \begin{pmatrix} * & * & & & & \\ * & * & * & & & \\ & * & * & * & & \\ & & * & * & * & \\ & & & * & * & * \\ & & & & * & * \end{pmatrix} \qquad \text{unreduced}$$

λ_1 distinct, all elements of v are non-zero
θ_1 distinct

Quasi-Sparsity for Tridiagonals

$$T = \begin{pmatrix} * & * & & & & \\ * & * & * & & & \\ & * & * & * & & \\ & & * & * & * & \\ & & & * & * & * \\ & & & & * & * \end{pmatrix}, \qquad v = \begin{pmatrix} v_1 \\ * \\ \mathbf{v_m} \\ \mathbf{v_{m+1}} \\ * \\ * \end{pmatrix}$$

$$\tau \equiv |\mathbf{v_m}|\,|\mathbf{v_{m+1}}| \Big/ \left\| \begin{pmatrix} v_1 \\ * \\ \mathbf{v_m} \end{pmatrix} \right\|^2$$

Quasi-Sparsity for Tridiagonals

- Determined by leading $m+1$ components of eigenvector
- Stronger than quasi-sparsity for dense matrices: $\tau \leq \rho$
- $\tau < \infty$ for unreduced tridiagonals

$\Longrightarrow$ Stronger Ritz value bounds for tridiagonals

Ritz Value Bounds for Tridiagonals

$$Tv = \lambda_1 v$$

$$T = \begin{pmatrix} * & * & & & & \\ * & * & * & & & \\ & * & * & \beta & & \\ & & \bar{\beta} & * & * & \\ & & & * & * & * \\ & & & & * & * \end{pmatrix}$$

Then

$$0 \le \theta_1 - \lambda_1 \le |\beta| \tau$$

Tightness of Ritz Value Bound

$$\theta_1 - \lambda_1 = c\,|\beta_m|\tau \qquad 0 \le c \le 1$$

How small can c be?

$$c \equiv \frac{\sum_{j=1}^{m} |\beta_j \ldots \beta_{m-1}\ \det(\lambda_1 I - T_{j-1})|^2}{|\det(\lambda_1 I - T_{m-1})|\ |\lambda_1 - \theta_2| \cdots |\lambda_1 - \theta_m|}$$

Example: Toeplitz Matrices

$$T \equiv \begin{pmatrix} \alpha & \beta & & \\ \beta & \alpha & \ddots & \\ & \ddots & \ddots & \beta \\ & & \beta & \alpha \end{pmatrix} \qquad \beta > 0$$

Exact error: $\theta_1 - \lambda_1 \approx \beta \left(\frac{\pi}{m+1}\right)^2$

Our bound: $\beta\tau \approx \beta \, \frac{6}{2m+1}$

Bound shows correct tendency: error decreases with β and $1/m$

Hermitian Tridiagonals

$Tv = \lambda_1 v$

If $|v_m|\,|v_{m+1}| \ll \left\| \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \right\|^2$ then

$$|\theta_1 - \lambda_1| \ll |\beta_m|$$

Smallest Ritz value $\approx$ smallest eigenvalue,

if mth and $(m+1)$st eigenvector components small

Relative Hermitian Bound

$$Hv = \lambda_1 v$$

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Then

$$\left| \frac{\theta_1 - \lambda_1}{\theta_1} \right| \leq \|H_{11}^{-1} H_{12}\| \, \rho \qquad \rho \equiv \frac{\|v_2\|}{\|v_1\|}$$

provided $\|H_{11}^{-1} H_{12}\| \, \rho < 1$

Simultaneous Bounds

Any m distinct eigenvalues of H: $\lambda_1 \ldots \lambda_m$

$$H = \begin{matrix} & m & \\ m & \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \end{matrix}$$

Eigenvalues of H_{11}: $\theta_1 \ldots \theta_m$

Eigenvectors $V \equiv \begin{pmatrix} v_1 & \ldots & v_m \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$

One-to-one matching of eigenvalues & Ritz values

(Partially) Hermitian Matrices

If H_{11} Hermitian and $\lambda_1 < \cdots < \lambda_m$ then

$$\left(\sum_{j=1}^{m} |\theta_j - \lambda_j|^2\right)^{1/2} \leq \sqrt{2}\,\|H_{12}\|_F \quad \underbrace{\|V_2 V_1^{-1}\|}_{\text{Quasi-sparsity}}$$

where $\theta_1 \leq \ldots \leq \theta_m$

Known ordering of eigenvalues & Ritz values

Real Ritz Values

If $\lambda_1 \ldots \lambda_m$ distinct and $\theta_1 < \cdots < \theta_m$ then

$$\left(\sum_{j=1}^{m} |\theta_j - \lambda_j|^2 \right)^{1/2} \leq \sqrt{2}\, \kappa(W)\, \|H_{12}\|_F \quad \underbrace{\|V_2 V_1^{-1}\|}_{\text{Quasi-sparsity}}$$

where $\Re(\lambda_1) \leq \ldots \leq \Re(\lambda_m)$

$\kappa(W) \equiv \|W\|\, \|W^{-1}\|$ where W Ritz vector matrix

$$H_{11} + E = V_1 \Lambda V_1^{-1} \qquad \|E\| \leq \|H_{12}\|\, \|V_2 V_1^{-1}\|$$

Advantages of Bound

Quasi-Sparsity $\|V_2 V_1^{-1}\|$

- No condition number of eigenvectors V
- If V quasi-sparse then H_{11} close to matrix with eigenvectors V_1 and eigenvalues λ_j
- Known ordering of eigenvalues & Ritz values
- Computable condition number $\kappa(W)$

General Matrices

If $\lambda_1 \ldots \lambda_m$ distinct and $\theta_1 \ldots \theta_m$ distinct then for some permutation $\sigma(\cdot)$

$$\left(\sum_{j=1}^{m} |\theta_{\sigma(j)} - \lambda_j|^2\right)^{1/2} \le \sqrt{m}\kappa(W)\|H_{12}\|_F \underbrace{\|V_2 V_1^{-1}\|}_{\text{Quasi-sparsity}}$$

One-to-one matching: eigenvalues $\leftrightarrow$ Ritz values

Real Eigenvalues

If $\lambda_1 < \cdots < \lambda_m$ then

$$\left(\sum_{j=1}^{m} |\theta_j - \lambda_j|^2\right)^{1/2} \leq \sqrt{2}\, \|H_{12}\|_F \underbrace{\|V_2\|\ \|V_1^{-1}\|}_{\text{Quasi-sparsity}}$$

where $\Re(\theta_1) \leq \cdots \leq \Re(\theta_m)$

No condition number for Ritz vectors
Known ordering: eigenvalues & Ritz values
But: Weaker quasi-sparsity

General Matrices

If $\lambda_1 \dots \lambda_m$ distinct then for some permutation $\sigma(\cdot)$

$$\left(\sum_{j=1}^{m} |\theta_{\sigma(j)} - \lambda_j|^2 \right)^{1/2} \le \sqrt{m}\, \|H_{12}\|_F \underbrace{\|V_2\|\ \|V_1^{-1}\|}_{\text{Quasi–sparsity}}$$

One-to-one matching: eigenvalues $\leftrightarrow$ Ritz values

Summary

- If eigenvector quasi-sparse then
  Ritz value $\approx$ eigenvalue
- Hermitian matrices:
  absolute & relative bounds for smallest eigenvalue
- General matrices:
  one-to-one matching bounds for several eigenvalues

  Only one eigenvector condition number