High-precision eigenvalue estimation with Kato-Lehmann-Goerisch's theorem

Xuefeng LIU

Tokyo Woman's Christian University (Japan)

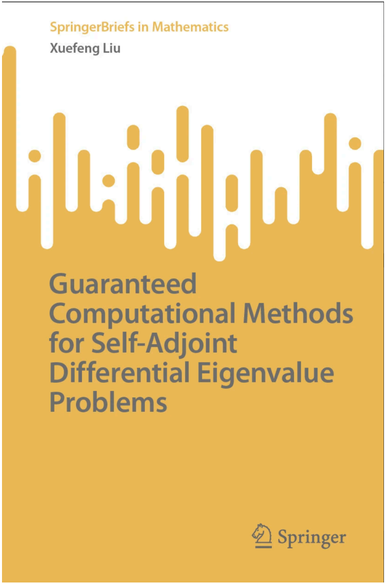
NMSP 2024 @Guiyang, China
August 5-9, 2024

(NMSP2019@Kyoto, NMSP2023@Hokkaido)

Outline

- Review of the projection based eigenvalue bounds
- Review of Kato's bound
- Lehmann–Goerisch's method
  - Comparison with Lehmann–Maehly's method
  - Comparison with Kato's bound

Note: The content in this slide is detailed my new book:
Guaranteed Computational Methods for Self-Adjoint Differential Eigenvalue Problems, (Springer).


SpringerBriefs in Mathematics
Xuefeng Liu
Guaranteed Computational Methods for Self-Adjoint Differential Eigenvalue Problems
Springer

July, 2024

Projection based eigenvalue bound and its defect

Model problem

Let Ω be a bounded domain and $V = H_0^1(\Omega)$.

The Laplacian eigenvalue problem

Find $u \in V$ and $\lambda \in R$ s.t.,

$$\int_\Omega \nabla u \cdot \nabla v \mathrm{d}x = \lambda \int_\Omega u \cdot v \mathrm{d}x \quad v \in V \ .$$

Let V^h be an approximation to V. V^h can be the conforming or non-conforming FEM spaces.

FEM approach to Laplacian eigenvalue problem

Find $u_h \in V^h$ and $\lambda_h \in R$ s.t.,

$$\int_\Omega \nabla u_h \cdot \nabla v_h \mathrm{d}x = \lambda_h \int_\Omega u_h \cdot v_h \mathrm{d}x \quad v \in V^h \ .$$

Projection error based eigenvalue bounds

Preparation

- $P_h : V \to V^h$: the projection that for $u \in V$,

$$(\nabla(u - P_h u), \nabla v_h) = 0, \quad \forall v_h \in V^h .$$

- A priori error estimation for P_h using quantity C_h:

$$\|u - P_h u\|_{L^2(\Omega)} \leq C_h \|\nabla(u - P_h u)\|_{L^2(\Omega)} .$$

Theorem for lower eigenvalue bounds

Lower eigenvalue bounds:

$$\frac{\lambda_{h,k}}{1 + C_h^2 \lambda_{h,k}} \leq \lambda_k \quad (k = 1, 2, \cdots, \mathrm{dim}(V^h)) .$$

Here, $\lambda_{h,k}$ is the approximation to target kth eigenvalue λ_k.

- For Crouzeix–Raviart FEM, we have $C_h = 0.1893h$.

[Birkhoff'1966] (cited as exercise of Spline Analysis of M. H. Schultz);
[Liu–Oishi'2013], [Liu'2015], [Carstensen–Gallistl'2014],
[Carstensen–Gedicke'2014]

Other variation of eigenvalue problem formulations

Eigenvalue problem: Find $\lambda \in R,\ u \in V$, s.t.,

$$a(u, v) = \lambda b(u, v) \ \ \forall v \in V \ .$$

Case 1:

$$\begin{cases} a(\cdot,\cdot): \text{ positive semi-definite, } \dim(\text{Ker}(a)) < \infty \\ b(\cdot,\cdot): \text{ positive definite} \end{cases}$$

For example, zero eigenvalues caused by Neumann boundary condition.

Case 2:

$$\begin{cases} a(\cdot,\cdot): \text{ positive definite;} \\ b(\cdot,\cdot): \text{ positive semi-definite} \end{cases}$$

For example, Steklov eigenvalue problem defined in $V = H^1(\Omega)$:

$$a(u, v) := \int_\Omega \nabla u \cdot \nabla v \ d\Omega, \ \ b(u, v) = \int_{\partial\Omega} uv \ ds$$

Application of projection based eigenvalue bounds

- Biharmonic operator: ([Liu-You, AMC'2018], [Liao-Shu-Liu, JJIAM'2019])

$$\Delta^2 u = -\lambda \Delta u \quad \text{or} \quad \Delta^2 u = \lambda u \quad \text{with boundary conditions}$$

- Steklov operator: ([Li-Liu, AM'2018], [You-Xie-Liu, SINUM'2019])

$$-\Delta u + cu = 0 \text{ in } \Omega, \quad \frac{\partial u}{\partial n} = \lambda u \text{ on } \partial\Omega$$

- Stokes operator: (2D: [Xie^2-Liu, JJIAM'2018], 3D: [Liu-You-Oishi-Nakao, JJIAM'2021])

$$-\Delta u + \nabla p = \lambda u, \ \text{div } u = 0 \quad \text{in } \Omega, \quad u = 0 \text{ on } \partial\Omega \ .$$

- Fluid–solid vibration problem [Y. Zhang-Y.D. Yang, CMA'2021]
- Maxwell operator: [D. Gallistl - V. Olkhovskiy, SINUM'2023]
- A survey of rigorous eigenvalue estimation: [Liu, JCAM'2020]
- Eigenvector estimation for clustered eigenvalues: [Liu–Vejchodsky, NM'2022, JCAM'2023] .

Advantages and defects

Advantages

1. The quantity C_h can be given explicit value. For example, for the Crouzeix–Raviart FEM on 2D domains, we have

$$C_h := 0.1893h \quad (C_h^2 := 0.036h^2) \quad (h : \text{mesh size})$$

For the Morley FEM on 2D domains, we have

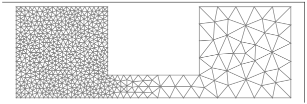
$$C_h^2 := 0.035h^2(\text{EVP} : \Delta^2 u = -\lambda \Delta u),\ C_h^2 := 0.006h^4(\text{EVP} : \Delta^2 u = \lambda u)$$

2. It works well for polygonal domains of arbitrary shapes.
3. Rather than the H^2 regularity, the H^1 regularity of the eigen-function is enough to derive eigenvalue bounds.

Advantages and defects (problems to solve)

Defects

1. The value of C_h is affected by the largest mesh element size. Thus, it is not suitable for locally refined meshes.
2. Generally, there is no high-order version of non-conforming FEMs. The hp-FEM cannot be utilized to provide efficient lower bounds.
3. The projection error based method must consider the worst case of the projection error.
   - When it is applied to the Steklov EVP, the eigenvalue bound has a sub-optimal convergence rates as $O(h)$ even for convex domains, compared to the convergence rate $O(h^2)$ of the approximation to leading eigenvalues.
4. For the Maxwell EVP, the estimation of C_h is very complicated even in 2D case. For 3D case, the lower bound, close to zero, is almost no-meaning.



Other approaches rather than projection based eigenvalue bounds

- Non-standard (or new) FEM method to provide direct eigenvalue bounds:
    - C Carstensen, S Puttkammer [SINUM'2023]
    - Carsten Carstensen, Qilong Zhai, Ran Zhang [SINUM'2020]
    - HHO method of [Carstensen– Ern–Puttkammer'2021]

Feature: Applicable to non-uniform meshes.

Re-examination of Kato's bound

Re-examination of Kato's bound

Take the eigenvalue problem of Δ as an example.

Kato's bound [Kato, 1949]

Let $\tilde{u} \in D(\Delta)$ be an approximate eigenvector, and $\tilde{\lambda} := \|\nabla \tilde{u}\|^2 / \|\tilde{u}\|^2$ and $\sigma := \| - \Delta \tilde{u} - \tilde{\lambda}\tilde{u}\| / \|\tilde{u}\|$. Suppose that μ and ν satisfy, for certain n,

$$\lambda_{n-1} \le \mu < \tilde{\lambda} < \nu \le \lambda_{n+1}.$$

Then,

$$\tilde{\lambda} - \frac{\sigma^2}{\nu - \tilde{\lambda}} \le \lambda_n \le \tilde{\lambda} + \frac{\sigma^2}{\tilde{\lambda} - \mu} .$$

∘ Kato, T., On the Upper and Lower Bounds of Eigenvalues. J. of Phys. Soci. of Japan, 1949

- A priori eigenvalue bounds μ and ν are needed;
- Quantities used in producing the eigenvalue bounds:

$$\|\Delta \tilde{u}\|, \|\nabla \tilde{u}\|, \|\tilde{u}\|.$$

How to relax the regularity requirement for $\tilde{u}$

Let us consider the Dirichlet eigenvalue problem.

Kato's bound or Lehmann's bound requires that $\tilde{u} \in D(-\Delta)$, since $\|-\Delta\tilde{u}\|$ is used to obtain bounds.

Goerisch's idea: Given $\tilde{v}$, choose $\tilde{u} \in H_0^1(\Omega)$

$$\tilde{u} := (-\Delta)^{-1}\tilde{v}, \quad \text{i.e.,} \;\; (\nabla\tilde{u}, \nabla g) = (\tilde{v}, g) \;\; \forall v \in H_0^1(\Omega) \;;$$

To avoid solving $\tilde{u}$ exactly, Goerisch choose $w \in H(\text{div})$ s.t.

$$(w, \nabla g) = (\tilde{v}, g) \;\; \forall v \in H_0^1(\Omega). \text{ That is, div } w + \tilde{v} = 0$$

Note that w is not unique, and for any candidate w, $\|w\| \geq \|\nabla\tilde{u}\|$.

Note : In Kato's bound, $\|\tilde{v}\|$, $\|\nabla\tilde{v}\|$, and $\|w\|$ (as upper bound of $\|\nabla\tilde{u}\|$) are enough to provide lower eigenvalue bounds.

Remark

Various results related to Kato's bounds

- The original result of Kato can also handle <u>clustered eigenvalues</u>, but it is not suitable in practical computation since it requires the <u>approximate eigenfunctions to be strictly orthogonal to each other</u>.
- Lehmann's theorem is almost the same as Kato's bound, but it can easily deal with <u>clustered eigenvalues</u>.
- Kato's bound or Lehmann's theorem requires that the approximate function $\hat{u}$ is smooth enough, while Goerisch's approach relaxes such a condition.

  For example, for the eigenvalue problem of $-\Delta$:

  - Kato's bound or Lehmann's theorem: $\hat{u} \in D(\Delta)$.
  - Lehmann-Goerisch's theorem: $\hat{v} \in H^1(\Omega)$ and $w \in H(\text{div};\Omega)$.

Lehmann-Goerisch's theorem

Assumptions and notation.

A1 D is a real vector space. M and N are symmetric bilinear forms on D; $M(f,f) > 0$ for all $f \in D$, $f \neq 0$.

A2 There exist sequences $\{\lambda_i\}_{i\in\mathbf{N}}$ and $\{\phi_i\}_{i\in\mathbf{N}}$ such that $\lambda_i \in \mathbf{R}$, $\phi_i \in D$, $M(\phi_i,\phi_k) = \delta_{ik}$ for $i,k \in \mathbf{N}$,

$$M(f,\phi_i) = \lambda_i N(f,\phi_k) \qquad \text{for all } f \in D, i \in \mathbf{N}. \tag{1}$$

$$N(f,f) = \sum_{i=1}^{\infty} (N(f,\phi_i))^2 \qquad \text{for all } f \in D, i \in \mathbf{N}. \tag{2}$$

A3 X is a real vector space; $T : D \to X$ is a linear operator; B is a symmetric bilinear form on X. $B(f,f) \geq 0$ for all $f \in X$ and $B(Tf,Tg) = M(f,g)$ for all $f,g \in D$.

A4 $n \in \mathbf{N}$, $v_i \in D$ for $i = 1,\cdots,n$. $w_i \in X$ satisfies

$$B(Tf,w_i) = N(f,v_i) \text{ for all } f \in D, i = 1,\cdots,n\,; \tag{3}$$

Lehmann-Goerisch's theorem

A5 $\rho \in \mathbf{R}$, $\rho > 0$. Define matrices as

$$A_0 := (M(v_i, v_k))_{i,k=1,\cdots,n}\,, \quad A_1 := (N(v_i, v_k))_{i,k=1,\cdots,n}\,,$$
$$A_2 := (B(w_i, w_k))_{i,k=1,\cdots,n}\,,$$
$$A^L = A_0 - \rho A_1, \quad B^L = A_0 - 2\rho A_1 + \rho^2 A_2;$$

B^L is positive definite. For $i = 1, \cdots, n$, the ith smallest eigenvalue of the eigenvalue problem $A^L z = \mu B^L z$ is denoted by μ_i.

Assertion: Suppose the ρ in A5 satisfies $\rho \leq \lambda_{m+1}$ and $-A^L$ is positivie definite. Then, a lower bound of λ_k $(1 \leq k \leq n)$ is given as

$$\rho - \rho/(1 - \mu_k) \leq \lambda_{m-k+1}\,.$$

Image of Lehmann–Goerisch's method

Given rough lower bound ρ of λ_{m+1}, with good approximation to eigenfunctions for $\lambda_{m-n+1}, \cdots, \lambda_m$, Lehmann–Goerisch's method provides sharp lower eigenvalue bounds.

$$\rho - \rho/(1-\mu_k) \leq \lambda_{m-k+1} \quad (1 \leq k \leq n)\,.$$

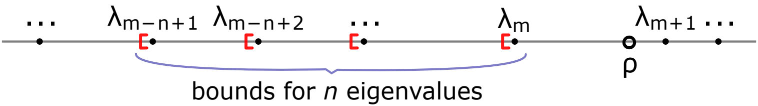


Figure: Eigenvalues to be involved in Lehmann–Goerisch's method.

Apply Lehmann-Goerisch's theorem to Laplacian

Settings

- Let $\Omega = (0,1)^2$. $D := H_0^1(\Omega)$, $X := (L_2(\Omega))^2$;
- $M(u,v) = (\nabla u, \nabla v)_{L_2(\Omega)}$, $N(u,v) = (u,v)_{L_2(\Omega)}$ for $u, v \in D$;

Eigenvalue problem of Laplacian

Find $\lambda \in R$ and $u \in D$ such that

$$M(u,v) = \lambda N(u,v) \quad \forall v \in D \,.$$

Rough eigenvalue bounds

Take the rough lower bound of the second eigenvalue by using projection based method (Liu'2015):

$$\rho = 48.0 < \lambda_2$$

Implementation of Lehmann–Goerisch's bound

Define $M(\cdot,\cdot)$, $b(\cdot,\cdot)$, T by

- $M(u,v) = (\nabla u, \nabla v)$ for $u, v \in D$, $\quad B(w,\tilde{w}) = (w,\tilde{w})$ for $w \in X$.
- $Tu := \nabla u$ for $u \in D$.

Selection of v_h and w_h:

- Take v_h as linear or quadratic conforming FEM solutions.
- Take w_h from H(div) FEM space, e.g., the Raviart–Thomas FEM space RT_h by solving the minimization problem:

$$\min_{w_h \in RT_h} \|w_h\|^2 \text{ subject to condition } \operatorname{div} w_h + v_h = 0 \quad (*)$$

Remark

Any w_h satisfying div $w_h + v_h = 0$ will produce eigenvalue bounds through Lehmann–Goerisch's method. It has been a problem for long time about how to select optimal w_h. Liu points out that the solution to $(*)$ gives the best selection of w_h, that is, the output eigenvalue bound will be the optimal one.

Computation results by Lehmann–Goerisch's bound

Computation results for uniform mesh with size as $h = 1/16$.

Linear FEM:

Upper bound (CG)	19.90871	(0.9%)
Lower bound (CR)	19.6098	(0.7%)
Lower bound (CG)	19.5393	(1.0%)
Lower bound (LG)	19.6214	(0.6%)

Quadratic FEM:

Upper bound (CG)	19.739467	(0.0013%)
Lower bound (CR)	-	-
Lower bound (CG)	19.647415	(0.47%)
Lower bound (LG)	19.739029	(0.0009%)

CR: Crouzeix–Raviart; CG: Conforming Galerkin; LG: Lehmann–Goerisch.

Computation example

Let us estimate the Dirichlet Laplacian eigenvalues over a dumbbell domain Ω. The approximate values of the eigenvalues are as follows,

$$\lambda_1 \approx \lambda_2 \approx 19.6755,\ \lambda_3 \approx \lambda_4 \approx 48.8407,\ \lambda_5 \approx \lambda_6 \approx 49.34761,\ \lambda_7 \approx 77.9216\ .$$

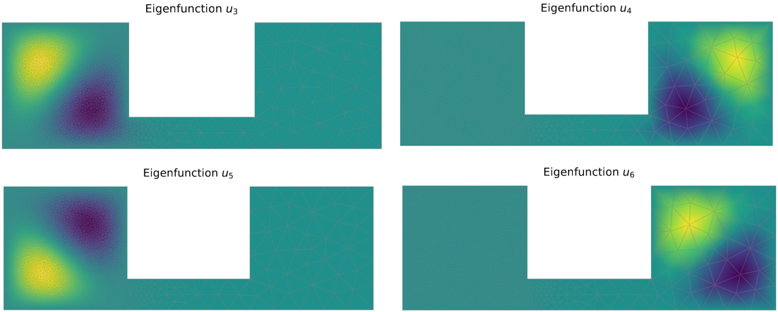


Figure: The eigenfunctions for the 4-6th eigenvalues (dumbbell)

Computation example

Local mesh refinement: the above initial mesh is partially refined to obtain the one below.

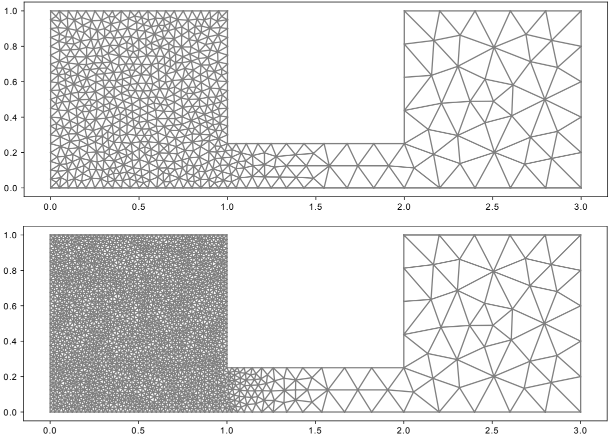


Figure: Local mesh refinement

Computation example

Table: Lower and upper bounds for the Laplacian eigenvalues (dumbbell)

Lower bounds by the Crouzeix–Raviart FEM

	initial mesh	refined mesh	rate
$\underline{\lambda}_3^{\mathsf{CR}}$	41.93548 (14.138%)	41.92631 (14.157%)	0.00
$\underline{\lambda}_4^{\mathsf{CR}}$	42.67912 (12.616%)	42.67894 (12.616%)	0.00
$\underline{\lambda}_5^{\mathsf{CR}}$	43.81675 (11.208%)	43.92367 (10.991%)	0.03
$\underline{\lambda}_6^{\mathsf{CR}}$	44.28652 (10.256%)	44.36634 (10.094%)	0.02

Lower bounds by Lehmann–Goerisch's method ($p = 2$)

	initial mesh	refined mesh	rate
$\underline{\lambda}_3^{\mathsf{LG}}$	48.45519 (0.789%)	48.45250 (0.795%)	0.01
$\underline{\lambda}_4^{\mathsf{LG}}$	48.80054 (0.082%)	48.81961 (0.043%)	0.93
$\underline{\lambda}_5^{\mathsf{LG}}$	49.09716 (0.508%)	49.09713 (0.508%)	0.00
$\underline{\lambda}_6^{\mathsf{LG}}$	49.34643 (0.002%)	49.34749 (0.0003%)	3.15

Note: $\lambda_3 = \lambda_4 \approx 48.8407,\ \lambda_5 = \lambda_6 \approx 49.34761.$

Remark for Lehmann–Goerisch's theorem

When Lehmann–Goerisch's theorem is applied to FEMs,

- It gives high-precision lower bounds as good as upper bounds.
- High-order FEM sheme.
- Graded mesh (non-uniform mesh).
- It works well for kth eigenvalue even for large k.

It can also be combined with the pectral method to have sharp eigenvalue bounds.

Conclusion

- The projection based eigenvalue bounds have several defects when applied to locally refined mesh and hp-FEM.
- Such defects can be completely improved by further combining the Lehmann–Goerisch method, which is an extension of Kato's bound.
- See application in Endo's talk in the afternoon.

Best solution for high-precesion eigenvalue bounds