Rayleigh Quotient Based Numerical Methods For Eigenvalue Problems

Ren-Cang Li

University of Texas at Arlington

Gene Golub SIAM Summer School 2013
10th Shanghai Summer School on Analysis and Numerics in Modern Sciences

July 22 – August 9, 2013

Overview

- Motivating Examples
- Hermitian Eigenvalue Problem – Basics
- Steep Descent/Ascent Type Methods
- Conjugate Gradient Type Methods
- Extending Min-Max Principles: Indefinite B
- Linear Response Eigenvalue Problem
- Quadratic Hyperbolic Eigenvalue Problem

Motivating Examples

- Density Functional Theory – Kohn-Sham Equation
- Data Mining – Trace Ratio Maximization

More in Chapter 10 of

Y. Saad. *Numerical Methods for Large Eigenvalue Problems*. SIAM, 2011.

Density Functional Theory (DFT)

Kohn-Sham Equation (Hohenberg and Kohn'64, Kohn and Sham'65):

$$\Big[-\frac{1}{2}\nabla^2 + \underbrace{\int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\, d\boldsymbol{r}' + \frac{\delta E_{\mathrm{xc}}(n(\boldsymbol{r}))}{\delta n(\boldsymbol{r})} + v_{\mathrm{ext}}(\boldsymbol{r})}_{v_{\mathrm{KS}}[n](\boldsymbol{r})} \Big] \phi_i(\boldsymbol{r}) = \lambda_i \phi_i(\boldsymbol{r}),$$

a remarkably successful theory to describe ground-state properties of condensed matter systems.

A nonlinear eigenvalue problem: Kohn-Sham (KS) operator depends on *electronic density* $n(\boldsymbol{r}) = \sum_{i=1}^{N_v} \phi_i(\boldsymbol{r})\phi_i^*(\boldsymbol{r})$ which depends on eigen-functions $\phi_i(\boldsymbol{r})$.

Usually solved by Self-Consistent-Field (SCF) iteration:

1) initial $n_0(\boldsymbol{r}) = \sum_{i=1}^{N_v} \phi_i^{(0)}(\boldsymbol{r})\phi_i^{(0)*}(\boldsymbol{r})$, and

2) repeat $\left[-\nabla^2/2 + v_{\mathrm{KS}}[n_j](\boldsymbol{r})\right] \phi_i^{(j+1)} = \lambda_i^{(j+1)} \phi_i^{(j+1)}(\boldsymbol{r})$.

Each inner-iteration is an eigenvalue problem.

Discretized Kohn-Sham Equation

Ways of discretizations: plane waves, finite differences, finite elements, localized orbitals, and wavelets.

Discretized Kohn-Sham Equation: $H(X)X = SX\Lambda$, $X^{\mathsf{H}}SX = I_{N_v}$.
$H(X)$ is symmetric, depends on X, eigenvalue matrix Λ is diagonal, and $S \succ 0$. Some discretizations: $S = I$.

Nonlinear eigenvalue problem, dependent on eigenvectors, as oppose to usually on the eigenvalues.

Usually solved by Self-Consistent-Field (SCF) iteration:

1) initial X_0, and
2) repeat $H(X_j)X_{j+1} = SX_{j+1}\Lambda_j$ for $j = 0, 1, \ldots$ until convergence.

Each inner-iteration is a symmetric eigenvalue problem.

References more accessible to numerical analysts:

Yousef Saad, James R. Chelikowsky, Suzanne M. Shontz, *Numerical Methods for Electronic Structure Calculations of Materials*, SIAM Rev. 52:1 (2010), 3-54.

C. Yang, J. C. Meza, B. Lee, and L.-W. Wang. KSSOLV—a MATLAB toolbox for solving the Kohn-Sham equations. *ACM Trans. Math. Software*, 36(2):1–35, 2009.

Trace Optimization

In Fisher linear discriminant analysis (LDA) for supervised learning, need to solve

$$\max_{V^{\mathsf{T}}V=I_k} \frac{\operatorname{trace}(V^{\mathsf{T}}AV)}{\operatorname{trace}(V^{\mathsf{T}}BV)},$$

where $A, B \in \mathbb{R}^{n\times n}$ symmetric, B positive semidefinite and $\operatorname{rank}(B) > n - k$.

$\operatorname{trace}(V^{\mathsf{T}}AV)$ represents the in-between scatter, while $\operatorname{trace}(V^{\mathsf{T}}BV)$ represents the within scatter. Maximizer V is used to project n-dimensional vectors (data) into k-dimensional vectors that best separates n-dimensional datasets into two or more datasets.

KKT condition for Maximizers:

$$\underbrace{\left[A - \frac{\operatorname{trace}(V^{\mathsf{T}}AV)}{\operatorname{trace}(V^{\mathsf{T}}BV)}\, B\right]}_{=:E(V)} V = V[V^{\mathsf{T}}E(V)\, V]$$

such that eigenvalues of $V^{\mathsf{T}}E(V)\, V$ are the k largest eigenvalues of $E(V)$.

Can be solved via SCF-like iteration; each inner iteration is a symmetric eigenvalue problem.

Trace Optimization

References for trace ratio problem:


T. Ngo, M. Bellalij, and Y. Saad. The trace ratio optimization problem for dimensionality reduction. *SIAM J. Matrix Anal. Appl.*, 31(5):2950–2971, 2010.


L.-H. Zhang, L.-Z. Liao, and M. K. Ng. Fast algorithms for the generalized Foley-Sammon discriminant analysis. *SIAM J. Matrix Anal. Appl.*, 31(4):1584–1605, 2010.

More eigenvalues arising from Data mining can be found in chapter 2 of


S. Yu, L.-C. Tranchevent, B. De Moor, and Y. Moreau. *Kernel-based Data Fusion for Machine Learning: Methods and Applications in Bioinformatics and Text Mining*. Springer, Berlin, 2011.

Basic Theory

- Hermitian $Ax = \lambda x$
- Hermitian $Ax = \lambda Bx$ $(B \succ 0)$
- Justifying Rayleigh-Ritz

Hermitian $Ax = \lambda x$

Hermitian $A = A^{\mathsf{H}} \in \mathbb{C}^{n\times n}$.

Eigenvalues λ_i and eigenvectors $u_i \in \mathbb{C}^n$.

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n, \quad u_i^{\mathsf{H}} u_j = \delta_{ij}, \quad Au_i = \lambda_i u_i.$$

Rich, elegant, and well-developed theories in "every" aspect ...

Popular References


R. Bhatia. *Matrix Analysis*. Springer, New York, 1996.


R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.


J. Demmel. *Applied Numerical Linear Algebra*. SIAM, Philadelphia, PA, 1997.


G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.


B. N. Parlett. *The Symmetric Eigenvalue Problem*. SIAM, Philadelphia, 1998.


Lloyd N. Trefethen and David Bau, III. *Numerical Linear Algebra*. SIAM, Philadelphia, 1997.


G. W. Stewart and Ji-Guang Sun. *Matrix Perturbation Theory*. Academic Press, Boston, 1990.

Courant-Fischer Theorem

Hermitian $A = A^{\mathsf{H}} \in \mathbb{C}^{n\times n}$. Eigenvalues: $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$.

Rayleigh quotient: $\rho(x) = \dfrac{x^{\mathsf{H}} A x}{x^{\mathsf{H}} x}$.

Courant (1920) and Fischer (1905)

$$\lambda_i = \min_{\dim \mathcal{X}=i} \max_{x\in\mathcal{X}} \rho(x), \quad \lambda_i = \max_{\operatorname{codim} \mathcal{X}=i-1} \min_{x\in\mathcal{X}} \rho(x).$$

In particular,

$$\lambda_1 = \min_x \rho(x), \quad \lambda_n = \max_x \rho(x). \tag{1}$$

- Can be used to justify Rayleigh-Ritz approximations for computational purposes.
- (1) is the foundation for using optimization techniques: steepest descent/ascent, CG type, for computing λ_1 and, with the help of deflation, other λ_j.

Trace Min/Trace Max

$A = A^{\mathsf{H}} \in \mathbb{C}^{n\times n}$. Eigenvalues: $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$.

Trace Min/Trace Max

$$\sum_{j=1}^{k} \lambda_j = \min_{X^{\mathsf{H}}X = I_k} \operatorname{trace}(X^{\mathsf{H}}AX),$$

$$\sum_{j=1}^{k} \lambda_{n-j+1} = \max_{X^{\mathsf{H}}X = I_k} \operatorname{trace}(X^{\mathsf{H}}AX).$$

- Can be used to justify Rayleigh-Ritz approximations for computational purposes.
- Rayleigh quotient matrix: $X^{\mathsf{H}}AX$, assuming $X^{\mathsf{H}}X = I_k$.

Cauchy Interlacing Theorem

Hermitian $A = A^{\mathsf{H}} \in \mathbb{C}^{n\times n}$. Eigenvalues: $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$.

$X \in \mathbb{C}^{n\times k}$, $k \le n$, $X^{\mathsf{H}}X = I_k$. Eigenvalue of $X^{\mathsf{H}}AX$: $\mu_1 \le \mu_2 \le \cdots \le \mu_k$.

Cauchy (1829)

$$\lambda_j \le \mu_j \le \lambda_{j+n-k} \quad \text{for } 1 \le j \le k.$$

Numerical implication: Pick X to "push" each μ_j down to λ_j or up to λ_{j+n-k}.

Hermitian $Ax = \lambda Bx$ $(B \succ 0)$

$A = A^{\mathsf{H}}, B = B^{\mathsf{H}} \in \mathbb{C}^{n\times n}$, and B positive definite.

Equivalency: $\quad Ax = \lambda Bx \quad \Leftrightarrow \quad \underbrace{B^{-1/2}AB^{-1/2}}_{=:\widehat{A}}\hat{x} = \lambda\hat{x},\ \hat{x} = B^{1/2}x.$

so same eigenvalues, and eigenvectors related by $\hat{x} = B^{1/2}x$.

Eigenvalues λ_i and eigenvectors $u_i \in \mathbb{C}^n$.

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n, \quad u_i^{\mathsf{H}} B u_j = \delta_{ij}, \quad Au_i = \lambda_i B u_i.$$

Rayleigh quotient: $\rho(x) := \dfrac{x^{\mathsf{H}}Ax}{x^{\mathsf{H}}Bx} \equiv \dfrac{\hat{x}^{\mathsf{H}}\widehat{A}\hat{x}}{\hat{x}^{\mathsf{H}}\hat{x}}.$

Verbatim translation of theoretical results for $\widehat{A}\hat{x} = \lambda\hat{x}$ to ones for $Ax = \lambda Bx$.

Hermitian $Ax = \lambda Bx$ $(B \succ 0)$

Courant (1920) and Fischer (1905)

$$\lambda_i = \min_{\dim \mathcal{X}=i} \max_{x\in\mathcal{X}} \rho(x), \quad \lambda_i = \max_{\operatorname{codim} \mathcal{X}=i-1} \min_{x\in\mathcal{X}} \rho(x).$$

In particular, $\lambda_1 = \min_x \rho(x)$, and $\lambda_n = \max_x \rho(x)$.

Trace Min/Trace Max

$$\sum_{j=1}^{k} \lambda_j = \min_{X^{\mathsf{H}}BX=I_k} \operatorname{trace}(X^{\mathsf{H}}AX), \quad \sum_{j=1}^{k} \lambda_{n-j+1} = \max_{X^{\mathsf{H}}BX=I_k} \operatorname{trace}(X^{\mathsf{H}}AX).$$

Cauchy (1829)

$X \in \mathbb{C}^{n\times k}$, $k \le n$, $\operatorname{rank}(X) = k$. Eigenvalues of $X^{\mathsf{H}}AX - \lambda X^{\mathsf{H}}BX$:
$\mu_1 \le \mu_2 \le \cdots \le \mu_k$.

$$\lambda_j \le \mu_j \le \lambda_{j+n-k} \quad \text{for } 1 \le j \le k.$$

Why Rayleigh-Ritz?

Two most important aspects in solving large scale eigenvalue problems:

1. building subspaces close to the desired eigenvectors (or invariant subspaces). E.g., Krylov subspaces.
2. seeking "*best possible*" approximations from the suitably built subspaces.

For 2nd aspect: given $\mathcal{Y} \in \mathbb{C}^n$ and $\dim \mathcal{Y} = m$, find the "*best possible*" approximations to some of the eigenvalues of $A - \lambda B$ using $\mathcal{Y}$.

Usually done by Rayleigh-Ritz Procedure. Let Y be $\mathcal{Y}$'s basis matrix.

Rayleigh-Ritz Procedure

1. Solve the eigenvalue problem for $Y^{\mathsf{H}}AY - \lambda Y^{\mathsf{H}}BY$: $Y^{\mathsf{H}}AYy_i = \mu_i Y^{\mathsf{H}}BYy_i$;
2. Approximate eigenvalues (Ritz values): $\mu_i (\approx \lambda_i)$;
   approximate eigenvectors (Ritz vectors): Yy_i.

But in what sense and why are those approximations "*best possible*"?

Why Rayleigh-Ritz? (cont'd)

Courant-Fischer: $\lambda_i = \min_{\dim \mathcal{X}=i} \max_{x\in\mathcal{X}} \rho(x)$ suggests that best possible approximation to λ_i should be taken as

$$\mu_i = \min_{\mathcal{X}\subset\mathcal{Y},\, \dim \mathcal{X}=i} \max_{x\in\mathcal{X}} \rho(x)$$

which is the ith eigenvalue of $Y^{\mathsf{H}}AY - \lambda Y^{\mathsf{H}}BY$.

Trace min principle: $\sum_{j=1}^{k} \lambda_j = \min_{X^{\mathsf{H}}BX=I_k} \operatorname{trace}(X^{\mathsf{H}}AX)$ suggests that best possible approximations to λ_i $(1 \le i \le k)$ should be gotten so that

$$\operatorname{trace}(X^{\mathsf{H}}AX) \text{ is minimized, subject to } \operatorname{span}(X) \subset \mathcal{Y},\ X^{\mathsf{H}}BX = I_k.$$

The optimal value is the sum of 1st k eigenvalues μ_i of $Y^{\mathsf{H}}AY - \lambda Y^{\mathsf{H}}BY$.
Consequently, $\mu_i \approx \lambda_i$ are "*best possible*".

Steepest Descent Methods

- Standard Steepest Descent Method
- Extended Steepest Descent Method
- Convergence Analysis
- Preconditioning Techniques
- Deflation

Problem: Hermitian $Ax = \lambda Bx$ ($B \succ 0$)

$A = A^{\mathsf{H}}, B = B^{\mathsf{H}} \in \mathbb{C}^{n \times n}$, and B positive definite.

Eigenvalues λ_i and eigenvectors $u_i \in \mathbb{C}^n$.

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n, \quad u_i^{\mathsf{H}} B u_j = \delta_{ij}, \quad Au_i = \lambda_i B u_i.$$

Rayleigh quotient: $\rho(x) := \dfrac{x^{\mathsf{H}} Ax}{x^{\mathsf{H}} Bx}$.

Interested in computing 1st eigenpair (λ_1, u_1). Later: Other eigenpairs with the help of deflation.

Largest eigenpairs: through considering $(-A) - \lambda B$ instead.

SD in general

SD method: a general technique to solve $\min f(x)$.

Steepest descent direction: at given x_0, along which direction p, f decreases fastest?

$$\min_p \left.\frac{d}{dt} f(x_0 + tp)\right|_{t=0} = \min_p p^{\mathsf{T}} \nabla f(x_0) = -\|\nabla f(x_0)\|_2 \tag{2}$$

optimal p is in the opposite direction of the gradient $\nabla f(x_0)$.

Plain SD: Given x_0, for $i = 0, 1, \ldots$ until convergence

$$t_i = \arg\min_t f(x_i + t\nabla f(x_i)), \quad x_{i+1} = x_i + t_i \nabla f(x_i). \tag{3}$$

Major work: solve $\min_t f(x_i + tp)$, so-called *line search*.

Food for thought. Derivation in (2) not quite right for real-valued function f of complex vector x.
In (3): $t \in \mathbb{R}$ or $t \in \mathbb{C}$ makes difference. $t \in \mathbb{C}$ potentially much more complicated!

Application to $\rho(x) = x^{\mathsf{H}}Ax/x^{\mathsf{H}}Bx$

Recall $\lambda_1 = \min_x \rho(x)$.

Gradient: $\nabla\rho(x) = \dfrac{2}{x^{\mathsf{H}}Bx}[Ax - \rho(x)Bx] =: \dfrac{2}{x^{\mathsf{H}}Bx}\, r(x)$. Note: $x^{\mathsf{H}}r(x) \equiv 0$.

$\|q\|$ tiny, up to 1st order:

$$\rho(x+q) = \frac{(x+q)^{\mathsf{H}}A(x+q)}{(x+q)^{\mathsf{H}}B(x+q)} = \frac{x^{\mathsf{H}}Ax + q^{\mathsf{H}}Ax + x^{\mathsf{H}}Aq}{x^{\mathsf{H}}Bx + q^{\mathsf{H}}Bx + x^{\mathsf{H}}Bq}$$
$$= \frac{x^{\mathsf{H}}Ax + q^{\mathsf{H}}Ax + x^{\mathsf{H}}Aq}{x^{\mathsf{H}}Bx} \cdot \left[1 - \frac{q^{\mathsf{H}}Bx + x^{\mathsf{H}}Bq}{x^{\mathsf{H}}Bx}\right] = \rho(x) + \frac{q^{\mathsf{H}}r(x) + r(x)^{\mathsf{H}}q}{x^{\mathsf{H}}Bx}.$$

Steepest descent direction: $\nabla\rho(x)$ parallel to *residual* $r(x) = A - \rho(x)Bx$.

Plain SD: Given x_0, for $i = 0, 1, \ldots$ until convergence

$$t_i = \arg\inf_t \rho(x_i + t\, r(x_i)), \quad x_{i+1} = x_i + t_i\, r(x_i).$$

- Major work: solve $\inf_t \rho(x_i + tp)$, so-called *line search*.
- When to stop?

Line Search $\inf_t \rho(x_i + t\,p)$

Can show $\inf_{t\in\mathbb{C}} \rho(x+tp) = \min_{|\xi|^2+|\zeta|^2>0} \rho(\xi x + \zeta p)$

which is smaller eigenvalue μ of 2×2 pencil $X^{\mathsf{H}}AX - \lambda X^{\mathsf{H}}BX$, where $X = [x, p]$.

Let $v = \begin{bmatrix}\nu_1\\ \nu_2\end{bmatrix}$ be the eigenvector. Then $\rho(Xv) = \mu$, and $Xv = \nu_1 x + \nu_2 p$. So

$$\arg\inf_{t\in\mathbb{C}} \rho(x+tp) =: t_{\text{opt}} = \begin{cases} \nu_2/\nu_1, & \text{if } \nu_1 \neq 0,\\ \infty, & \text{if } \nu_1 = 0.\end{cases}$$

Interpret $t_{\text{opt}} = \infty$ in the sense $\lim_{t\to\infty} \rho(x+tp) = \rho(p)$.

$$\rho(y) = \inf_{t\in\mathbb{C}} \rho(x+tp), \quad y = \begin{cases} x + t_{\text{opt}}\,p & \text{if } t_{\text{opt}} \text{ is finite,}\\ p & \text{otherwise}\end{cases}$$

A Theorem for Line Search

Line Search

Suppose x, p are linearly independent. Then $x^{\mathsf{H}} r(y) = 0$ and $p^{\mathsf{H}} r(y) = 0$.

Proof

$p^{\mathsf{H}} r(y) = 0$: True if $y = p$, i.e., $t_{\text{opt}} = \infty$. Otherwise

$$y = x + t_{\text{opt}} p, \quad \rho(y) = \min_{t \in \mathbb{C}} \rho(x + tp) = \min_{s \in \mathbb{C}} \rho(y + sp).$$

Optimal at $s = 0$. $\rho(y + sp) = \rho(y) + \frac{2}{y^{\mathsf{H}} B y} \Re(s p^{\mathsf{H}} r(y)) + \mathcal{O}(s^2)$ implies $p^{\mathsf{H}} r(y) = 0$.

$x^{\mathsf{H}} r(y) = 0$: True if $y = x$, $t_{\text{opt}} = 0$. and thus $x^{\mathsf{H}} r(y) = x^{\mathsf{H}} r(x) = 0$. Otherwise

$$y = \arg\inf_{s \in \mathbb{C}} \rho(p + sx).$$

Therefore $x^{\mathsf{H}} r(y) = 0$.

Stopping Criteria

Common one: check if $\|r(x)\|$ tiny enough. Reason: Easy to use and available.

$$\text{if} \quad \frac{\|r(\boldsymbol{x})\|_2}{\|A\boldsymbol{x}\|_2 + |\rho(\boldsymbol{x})| \, \|B\boldsymbol{x}\|_2} \leq \texttt{rtol}.$$

Implication: $(\rho(\boldsymbol{x}), \boldsymbol{x})$ is an exact eigenpair of $(A + E) - \lambda B$ for some Hermitian matrix E.

Can prove that (suppose $\|\boldsymbol{x}\|_2 = 1$)

$$\min \|E\|_2 = \|r(\boldsymbol{x})\|_2, \quad \min \|E\|_F = \sqrt{2}\|r(\boldsymbol{x})\|_2.$$

More can be found in Chapter 5 of:

Zhaojun Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst (editors). *Templates for the solution of Algebraic Eigenvalue Problems: A Practical Guide*. SIAM, Philadelphia, 2000.

Framework of SD

Steepest Descent method

Given an initial approximation $\boldsymbol{x}_0$ to u_1, and a relative tolerance `rtol`, the algorithm attempts to compute an approximate eigenpair to (λ_1, u_1) with the prescribed `rtol`.

1: $\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|_B$, $\boldsymbol{\rho}_0 = \boldsymbol{x}_0^{\mathsf{H}} A\boldsymbol{x}_0$, $\boldsymbol{r}_0 = A\boldsymbol{x}_0 - \boldsymbol{\rho}_0 B\boldsymbol{x}_0$;
2: **for** $\ell = 0, 1, \ldots$ **do**
3: **if** $\|\boldsymbol{r}_\ell\|_2/(\|A\boldsymbol{x}_\ell\|_2 + |\boldsymbol{\rho}_\ell|\,\|B\boldsymbol{x}_\ell\|_2) \leq$ `rtol` **then**
4: **BREAK**;
5: **else**
6: compute the smaller eigenvalue μ and corresponding eigenvector v of $Z^{\mathsf{H}}(A - \lambda B)Z$, where $Z = [\boldsymbol{x}_\ell, \boldsymbol{r}_\ell]$;
7: $\hat{x} = Zv$, $\boldsymbol{x}_{\ell+1} = \hat{x}/\|\hat{x}\|_B$;
8: $\boldsymbol{\rho}_{\ell+1} = \mu$, $\boldsymbol{r}_{\ell+1} = A\boldsymbol{x}_{\ell+1} - \boldsymbol{\rho}_{\ell+1} B\boldsymbol{x}_{\ell+1}$;
9: **end if**
10: **end for**
11: **return** $(\boldsymbol{\rho}_\ell, \boldsymbol{x}_\ell)$ as an approximate eigenpair to (λ_1, u_1).

Note: At Line 6, $\operatorname{rank}(Z) = 2$ always unless $\boldsymbol{r}_\ell = 0$ because $\boldsymbol{x}_\ell^{\mathsf{H}} \boldsymbol{r}_\ell = 0$.

SD: Pros and Cons

Pros: Easy to implement; low memory requirement.
Cons: Possibly slow to converge, sometimes unbearably slow.

Well-known: SD slowly moves in zigzag towards an optimal point when the contours near the point are extremely flat.

Ways to rescue:

- Extended the search space: "line search" to "subspace search"
- Modify the search direction: move away from the steepest descent direction $-\nabla\rho(x)$
- Combination

Extended SD Method

Seek to extend the search space *naturally*.

SD search space $\text{span}\{x, r(x)\}$. Note $r(x) = Ax - \rho(x)Bx = [A - \rho(x)B]x$.

$$\text{span}\{x, r(x)\} = \text{span}\{x, [A - \rho(x)B]x\} = \mathcal{K}_2([A - \rho(x)B], x)$$

the *2nd Krylov subspace* of $A - \rho(x)B$ on x.

Naturally extend $\mathcal{K}_2([A - \rho(x)B], x)$ to

$$\mathcal{K}_m([A - \rho(x)B], x) = \text{span}\{x, [A - \rho(x)B]x, \ldots, [A - \rho(x)B]^{m-1}x\},$$

the *mth Krylov subspace* of $A - \rho(x)B$ on x.

Call resulting method *extended steepest descent method* (ESD). It is in fact the inverse free Krylov subspace method of Golub and Ye (2002).

Framework of ESD

Extended Steepest Descent method

Given an initial approximation $\boldsymbol{x}_0$ to u_1, a relative tolerance `rtol`, and an integer $m \geq 2$, the algorithm attempts to compute an approximate eigenpair to (λ_1, u_1) with the prescribed `rtol`.

1: $\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|_B$, $\boldsymbol{\rho}_0 = \boldsymbol{x}_0^{\mathsf{H}} A\boldsymbol{x}_0$, $\boldsymbol{r}_0 = A\boldsymbol{x}_0 - \boldsymbol{\rho}_0 B\boldsymbol{x}_0$;
2: **for** $\ell = 0, 1, \ldots$ **do**
3: **if** $\|\boldsymbol{r}_\ell\|_2/(\|A\boldsymbol{x}_\ell\|_2 + |\boldsymbol{\rho}_\ell|\,\|B\boldsymbol{x}_\ell\|_2) \leq$ `rtol` **then**
4: **BREAK**;
5: **else**
6: compute a basis matrix $Z \in \mathbb{C}^{n\times m}$ of Krylov subspace $\mathcal{K}_m(A - \boldsymbol{\rho}_\ell B, \boldsymbol{x}_\ell)$;
7: compute the smallest eigenvalue μ and corresponding eigenvector v of $Z^{\mathsf{H}}(A - \lambda B)Z$;
8: $y = Zv$, $\boldsymbol{x}_{\ell+1} = y/\|y\|_B$;
9: $\boldsymbol{\rho}_{\ell+1} = \mu$, $\boldsymbol{r}_{\ell+1} = A\boldsymbol{x}_{\ell+1} - \boldsymbol{\rho}_{\ell+1} B\boldsymbol{x}_{\ell+1}$;
10: **end if**
11: **end for**
12: **return** $(\boldsymbol{\rho}_\ell, \boldsymbol{x}_\ell)$ as an approximate eigenpair to (λ_1, u_1).

Note: If $B = I$, it is equivalent to *Restarted Lanczos*

Basis of $\mathcal{K}_m(A - \rho B, x)$

$C = A - \rho B$ is Hermitian.

Lanczos process

1: $z_1 = x/\|x\|_2$, $\beta_0 = 0$; $z_0 = 0$;
2: **for** $j = 1, 2, \ldots, k$ **do**
3: $\quad z = Cz_j$, $\alpha_j = z_j^{\mathrm{H}} z$;
4: $\quad z = z - \alpha_j z_j - \beta_{j-1} z_{j-1}$, $\beta_j = \|z\|_2$;
5: $\quad$ **if** $\beta_j = 0$ **then**
6: $\quad\quad$ **BREAK**;
7: $\quad$ **else**
8: $\quad\quad z_{j+1} = z/\beta_j$;
9: $\quad$ **end if**
10: **end for**

- Keep Az_j and Bz_j for projecting A and B later. Or, just z_j but solve $Z^{\mathrm{H}} C Z - \lambda Z^{\mathrm{H}} B Z$ instead
- Implemented as is, $Z = [z_1, \ldots, z_m]$ may lose orthogonality — partial or full re-orthogonalization should be used. Pose little problem since usually m is modest.
- Possibly $\dim \mathcal{K}_m(A - \rho B, x) < m$. Pose no problem — Z has fewer than m columns

Global Convergence

Convergence

For SD and ESD, $\boldsymbol{\rho}_\ell$ converges to some eigenvalue $\hat{\lambda}$ of $A - \lambda B$ and $\|(A - \hat{\lambda}B)\boldsymbol{x}_\ell\|_2 \to 0$.

Proof.

1) $\{\boldsymbol{\rho}_\ell\}$ monotonically decreasing and $\boldsymbol{\rho}_\ell \geq \lambda_1 \quad \Rightarrow \quad \boldsymbol{\rho}_\ell \to \hat{\lambda}$.
2) $\{\boldsymbol{x}_\ell\}$ bounded in $\mathbb{C}^n \quad \Rightarrow \quad$ convergent $\{\boldsymbol{x}_{n_\ell}\}$, $\boldsymbol{x}_{n_\ell} \to \hat{x}$.
3) $\boldsymbol{x}_{n_\ell}^{\mathsf{H}}(A - \boldsymbol{\rho}_{n_\ell}B)\boldsymbol{x}_{n_\ell} = 0 \quad \Rightarrow \quad \hat{x}^{\mathsf{H}}\hat{r} = \hat{x}^{\mathsf{H}}(A - \hat{\lambda}B)\hat{x} = 0$.
4) Claim $\hat{r} = 0$. Otherwise $\hat{r} \neq 0$, $\operatorname{rank}([\hat{x}, \hat{r}]) = 2$, and

$$\widehat{A} - \hat{\lambda}\widehat{B} := \begin{bmatrix} \hat{x}^{\mathsf{H}} \\ \hat{r}^{\mathsf{H}} \end{bmatrix} A[\hat{x}, \hat{r}] - \hat{\lambda} \begin{bmatrix} \hat{x}^{\mathsf{H}} \\ \hat{r}^{\mathsf{H}} \end{bmatrix} B[\hat{x}, \hat{r}] = \begin{bmatrix} 0 & \hat{r}^{\mathsf{H}}\hat{r} \\ \hat{r}^{\mathsf{H}}\hat{r} & \hat{r}^{\mathsf{H}}(A - \hat{\lambda}B)\hat{r} \end{bmatrix} \text{ is indefinite.}$$

Smaller eigenvalue μ of $\widehat{A} - \lambda\widehat{B}$: $\mu < \hat{\lambda}$. Let

$$\widehat{A}_\ell = [\boldsymbol{x}_\ell, \boldsymbol{r}_\ell]^{\mathsf{H}} A [\boldsymbol{x}_\ell, \boldsymbol{r}_\ell], \quad \widehat{B}_\ell = [\boldsymbol{x}_\ell, \boldsymbol{r}_\ell]^{\mathsf{H}} B [\boldsymbol{x}_\ell, \boldsymbol{r}_\ell],$$

$\mu_{\ell+1}$ smaller eigenvalue of $\widehat{A}_\ell - \lambda\widehat{B}_\ell$. Then

(i) $\widehat{A}_{n_\ell} \to \widehat{A}, \quad \widehat{B}_{n_\ell} \to \widehat{B} \quad \Rightarrow \quad \mu_{n_\ell+1} \to \mu$,

(ii) $\rho_{n_\ell+1} \leq \mu_{n_\ell+1} \quad \Rightarrow \quad \hat{\lambda} = \lim_{i\to\infty} \rho_{n_\ell+1} \leq \lim_{i\to\infty} \mu_{n_\ell+1} = \mu$,

a contradiction! So $\hat{r} = 0$. □

Digression: Chebyshev Polynomial

The jth Chebyshev polynomial of the first kind $\mathscr{T}_j(t)$

$$\begin{aligned}\mathscr{T}_j(t) &= \cos(j \arccos t) && \text{for } |t| \le 1,\\ &= \frac{1}{2}\left[\left(t+\sqrt{t^2-1}\right)^j + \left(t+\sqrt{t^2-1}\right)^{-j}\right] && \text{for } t \ge 1.\end{aligned}$$

Or, $\mathscr{T}_0(t) = 1$, $\mathscr{T}_1(t) = t$, and $\mathscr{T}_j(t) = 2\mathscr{T}_{j-1}(t) - \mathscr{T}_{j-2}(t)$ for $j \ge 2$.
Numerous optimal properties among polynomials

- $\deg(p) \le j$, $|p(t)| \le 1$ for $t \in [-1,1] \Rightarrow |p(t)| \le |\mathscr{T}_j(t)|$ for $t \notin [-1,1]$.

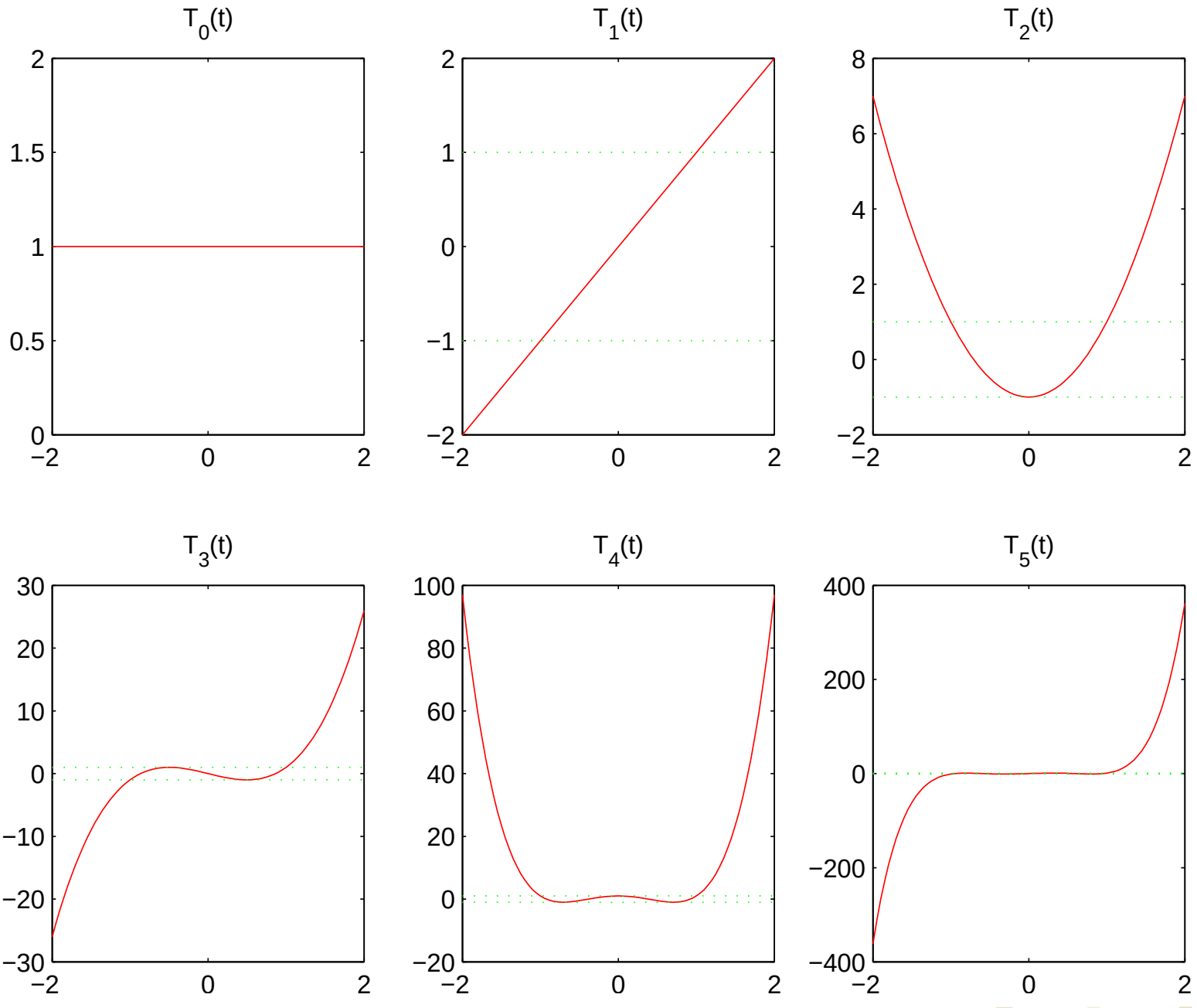
I.e., $|\mathscr{T}_j(t)| \le 1$ for $|t| \le 1$ and $|\mathscr{T}_j(t)|$ grows fastest. (Sample plots next slide.)

$$\left|\mathscr{T}_j\left(\frac{1+t}{1-t}\right)\right| = \left|\mathscr{T}_j\left(\frac{t+1}{t-1}\right)\right| = \frac{1}{2}\left[\Delta_t^j + \Delta_t^{-j}\right] \quad \text{for } 1 \ne t > 0,$$

where $\Delta_t := \dfrac{\sqrt{t}+1}{|\sqrt{t}-1|}$ for $t > 0$.

Frequently show up in numerical analysis and computations: Chebyshev acceleration in iterative methods, convergence of CG and Lanczos methods.

Chebyshev Polynomial (sample plots)



Chebyshev Polynomial: typical use

Problem Given $[\alpha, \beta]$ and $\gamma \notin [\alpha, \beta]$, seek a polynomial p with $\deg(p) \leq m$ such that $p(\gamma) = 1$ and $\max_{x \in [\alpha,\beta]} |p(x)|$ is minimized.

Define 1-1 mapping $x \in [\alpha, \beta] \rightarrow t \equiv t(x) := \frac{2}{\beta-\alpha}\left(x - \frac{\alpha+\beta}{2}\right) \in [-1, 1]$,

$$t(\gamma) = -\frac{1 + \frac{\alpha-\gamma}{\beta-\gamma}}{1 - \frac{\alpha-\gamma}{\beta-\gamma}} \quad \text{for } \gamma < \alpha, \text{ and} \quad \frac{1 + \frac{\gamma-\beta}{\gamma-\alpha}}{1 - \frac{\gamma-\beta}{\gamma-\alpha}} \quad \text{for } \beta < \gamma.$$

Optimal $p(x) = \dfrac{\mathscr{T}_m(t(x))}{\mathscr{T}_m(t(\gamma))}$:

$$p(\gamma) = 1, \quad \max_{x \in [\alpha,\beta]} |p(x)| = \frac{1}{|\mathscr{T}_m(t(\gamma))|} = 2\left[\Delta_\eta^j + \Delta_\eta^{-j}\right]^{-1},$$

$$\Delta_\eta = \frac{1 + \sqrt{\eta}}{1 - \sqrt{\eta}}, \quad \eta = \frac{\alpha - \gamma}{\beta - \gamma} \quad \text{for } \gamma < \alpha, \text{ and} \quad \eta = \frac{\gamma - \beta}{\gamma - \alpha} \quad \text{for } \beta < \gamma.$$

Convergence Rate

Theorem on Convergence Rate (Golub & Ye, 2002)

Suppose λ_1 is simple, i.e., $\lambda_1 < \lambda_2$, and $\lambda_1 < \boldsymbol{\rho}_\ell < \lambda_2$. Let

$$\omega_1 < \omega_2 \leq \cdots \leq \omega_n$$

be the eigenvalues of $A - \boldsymbol{\rho}_\ell B$ and v_1 be an eigenvector corresponding to ω_1. Then

$$\boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq (\boldsymbol{\rho}_\ell - \lambda_1)\epsilon_m^2 + 2(\boldsymbol{\rho}_\ell - \lambda_1)^{3/2}\epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + \delta_\ell$$

where

$$0 \leq \delta_\ell := \boldsymbol{\rho}_\ell - \lambda_1 + \omega_1 \frac{v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} = \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2),$$

$$\epsilon_m := \min_{f \in \mathbb{P}_{m-1}, f(\omega_1)=1} \max_{j>1} |f(\omega_j)|.$$

Discussion

$\epsilon_m := \min_{f\in\mathbb{P}_{m-1}, f(\omega_1)=1} \max_{j>1} |f(\omega_j)|$ usually unknown, except,

1. $m = 2$ for which the optimal f_{opt} is

$$f_{\text{opt}}(t) = \frac{t - (\omega_2+\omega_n)/2}{\omega_1 - (\omega_2+\omega_n)/2} \in \mathbb{P}_{m-1},\ f_{\text{opt}}(\omega_1) = 1,\ \max_{j>1} |f_{\text{opt}}(\omega_j)| = |f_{\text{opt}}(\omega_2)| < 1.,$$

$$\epsilon_2 = \frac{1-\eta}{1+\eta}, \quad \eta = \frac{\omega_2 - \omega_1}{\omega_n - \omega_1}.$$

2. $\omega_2 = \cdots = \omega_n$ for which $f_{\text{opt}}(t) = (t-\omega_2)/(\omega_1 - \omega_2)$ and $\epsilon_m = 0$ for all $m \geq 2$.

In general ϵ_m can be bounded by using the Chebyshev polynomial

$$f(t) = \mathscr{T}_{m-1}\left(\frac{2t - (\omega_n+\omega_2)}{\omega_n - \omega_2}\right) \Big/ \mathscr{T}_{m-1}\left(\frac{1+\eta}{1-\eta}\right), \qquad f(\omega_1) = 1,$$

$$\epsilon_m \leq \max_{\omega_2 \leq t \omega_n} |f(t)| = \left[\mathscr{T}_{m-1}\left(\frac{1+\eta}{1-\eta}\right)\right]^{-1} = 2\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-1}.$$

Discussion (cont'd)

$$\boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq (\boldsymbol{\rho}_\ell - \lambda_1)\epsilon_m^2 + 2(\boldsymbol{\rho}_\ell - \lambda_1)^{3/2}\epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2).$$

Ignoring high order terms, $\dfrac{\boldsymbol{\rho}_{\ell+1} - \lambda_1}{\boldsymbol{\rho}_\ell - \lambda_1} \lessapprox \epsilon_m^2$.

If $\epsilon_m = 0$ (unlikely, however), then $\boldsymbol{\rho}_{\ell+1} - \lambda_1 = \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2)$, quadratic convergence.

Locally, $\boldsymbol{\rho}_\ell \approx \lambda_1$, $\text{eig}(A - \boldsymbol{\rho}_\ell B) \approx \text{eig}(A - \lambda_1 B) = \{0 = \gamma_1 < \gamma_2 \leq \cdots \leq \gamma_n\}$. So

$$\epsilon_m \approx \min_{f \in \mathbb{P}_{m-1}, f(\gamma_1)=1} \max_{j>1} |f(\gamma_j)| \leq 2\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-1},$$

$$\Delta_\eta = \frac{1+\sqrt{\eta}}{1-\sqrt{\eta}}, \quad \eta = \frac{\gamma_2 - \gamma_1}{\gamma_n - \gamma_1}.$$

Observation. ϵ_m depends on $\text{eig}(A - \boldsymbol{\rho}_\ell B)$, not on $\text{eig}(A, B)$. This is the Key for preconditioning later: transforming $A - \lambda B$ to preserve $\text{eig}(A, B)$ but make $\text{eig}(A - \boldsymbol{\rho}_\ell B)$ more preferable.

Key Lemma

Lemma (Golub & Ye, 2002)

Let (ω_1, v_1) be the smallest eigenvalue of $A - \boldsymbol{\rho}_\ell B$, i.e.. Then

$$-\omega_1 \frac{v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} \leq \boldsymbol{\rho}_\ell - \lambda_1 \leq -\omega_1 \frac{u_1^{\mathsf{H}} u_1}{u_1^{\mathsf{H}} B u_1}. \tag{4}$$

Asymptotically, if λ_1 is a simple eigenvalue of $A - \lambda B$, then as $\boldsymbol{\rho}_\ell \to \lambda_1$,

$$\omega_1 \frac{v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} = (\lambda_1 - \boldsymbol{\rho}_\ell) + \mathcal{O}(|\lambda_1 - \boldsymbol{\rho}_\ell|^2). \tag{5}$$

Importance. Relate $\lambda_1 - \boldsymbol{\rho}_\ell$ to ω_1. Special case: $B = I$, $\omega_1 = \lambda_1 - \boldsymbol{\rho}_\ell$.

Proof of Key Lemma

1) $\boldsymbol{\rho}_\ell \geq \lambda_1$ always. If $\boldsymbol{\rho}_\ell = \lambda_1$, then $\omega_1 = 0$. No proof needed.
2) Suppose $\boldsymbol{\rho}_\ell > \lambda_1$. $A - \boldsymbol{\rho}_\ell B$ is indefinite and hence $\omega_1 < 0$. We have

$$(A - \boldsymbol{\rho}_\ell B)v_1 = \omega_1 v_1, \quad (A - \omega_1 I - \boldsymbol{\rho}_\ell B)v_1 = 0, \quad A - \omega_1 I - \boldsymbol{\rho}_\ell B \succeq 0.$$

Therefore $(\boldsymbol{\rho}_\ell, v_1)$ is the smallest eigenpair of $(A - \omega_1 I) - \lambda B$.
3) Note also (λ_1, u_1) is the smallest eigenpair of $A - \lambda B$.

$$\begin{aligned}
\boldsymbol{\rho}_\ell &= \frac{v_1^{\mathsf{H}}(A - \omega_1 I)v_1}{v_1^{\mathsf{H}} B v_1} = \frac{v_1^{\mathsf{H}} A v_1}{v_1^{\mathsf{H}} B v_1} + \frac{-\omega_1 v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} \\
&\geq \min_x \frac{x^{\mathsf{H}} A x}{x^{\mathsf{H}} B x} + \frac{-\omega_1 v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} = \lambda_1 + \frac{-\omega_1 v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1}, \\
\lambda_1 &= \frac{u_1^{\mathsf{H}} A u_1}{u_1^{\mathsf{H}} B u_1} = \frac{u_1^{\mathsf{H}}(A - \boldsymbol{\rho}_\ell B)u_1}{u_1^{\mathsf{H}} u_1} \cdot \frac{u_1^{\mathsf{H}} u_1}{u_1^{\mathsf{H}} B u_1} + \boldsymbol{\rho}_\ell \\
&\geq \min_x \frac{x^{\mathsf{H}}(A - \boldsymbol{\rho}_\ell B)x}{x^{\mathsf{H}} x} \cdot \frac{u_1^{\mathsf{H}} u_1}{u_1^{\mathsf{H}} B u_1} + \boldsymbol{\rho}_\ell = \omega_1 \cdot \frac{u_1^{\mathsf{H}} u_1}{u_1^{\mathsf{H}} B u_1} + \boldsymbol{\rho}_\ell.
\end{aligned}$$

Together yielding $-\omega_1 \dfrac{v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} \leq \boldsymbol{\rho}_\ell - \lambda_1 \leq -\omega_1 \dfrac{u_1^{\mathsf{H}} u_1}{u_1^{\mathsf{H}} B u_1}$.

Proof of Key Lemma

4) $\omega_1(t) = \lambda_{\min}(A - tB)$, for t near λ_1. Then $\omega_1(\lambda_1) = 0$ and $\omega_1(\boldsymbol{\rho}_\ell) = \omega_1$.
5) $\omega_1(\lambda_1) = 0$ is a simple eigenvalue of $A - \lambda_1 B$. So $\omega_1(t)$ is differentiable in a neighborhood of λ_1.
6) Expand $\omega_1(t)$ at $\boldsymbol{\rho}_\ell$, sufficiently close to λ_1. Can prove $\omega_1'(\boldsymbol{\rho}_\ell) = -\frac{v_1^{\mathsf{H}} B v_1}{v_1^{\mathsf{H}} v_1}$. Hence

$$\begin{aligned}\omega_1(t) &= \omega_1(\boldsymbol{\rho}_\ell) + \sigma_1'(\boldsymbol{\rho}_\ell)(t - \boldsymbol{\rho}_\ell) + \mathcal{O}(|t - \boldsymbol{\rho}_\ell|^2) \\ &= \omega_1 - \frac{v_1^{\mathsf{H}} B v_1}{v_1^{\mathsf{H}} v_1}(t - \boldsymbol{\rho}_\ell) + \mathcal{O}(|t - \boldsymbol{\rho}_\ell|^2).\end{aligned}$$

Setting $t = \lambda_1$,

$$0 = \omega_1(\lambda_1) = \omega_1 - \frac{v_1^{\mathsf{H}} B v_1}{v_1^{\mathsf{H}} v_1}(\lambda_1 - \boldsymbol{\rho}_\ell) + \mathcal{O}(|\lambda_1 - \boldsymbol{\rho}_\ell|^2),$$

from which $$\omega_1 \frac{v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} = (\lambda_1 - \boldsymbol{\rho}_\ell) + \mathcal{O}(|\lambda_1 - \boldsymbol{\rho}_\ell|^2).$$ □

Convergence Rate

Theorem on Convergence Rate (Golub & Ye, 2002)

Suppose λ_1 is simple, i.e., $\lambda_1 < \lambda_2$, and $\lambda_1 < \boldsymbol{\rho}_\ell < \lambda_2$. Let

$$\omega_1 < \omega_2 \leq \cdots \leq \omega_n$$

be the eigenvalues of $A - \boldsymbol{\rho}_\ell B$ and v_1 be an eigenvector corresponding to ω_1. Then

$$\boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq (\boldsymbol{\rho}_\ell - \lambda_1)\epsilon_m^2 + 2(\boldsymbol{\rho}_\ell - \lambda_1)^{3/2}\epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + \delta_\ell$$

where

$$0 \leq \delta_\ell := \boldsymbol{\rho}_\ell - \lambda_1 + \omega_1 \frac{v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} = \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2),$$

$$\epsilon_m := \min_{f \in \mathbb{P}_{m-1}, f(\omega_1)=1} \max_{j>1} |f(\omega_j)|.$$

Proof

1) $C = A - \rho_\ell B = V\Omega V^{\mathsf{H}}$ (eigen-decomposition), $V = [v_1, v_2, \cdots, v_n]$ (orthogonal), $\Omega = \mathrm{diag}(\omega_1, \omega_2, \cdots, \omega_n)$.
2) $\mathcal{K}_m \equiv \mathcal{K}_m(C, \boldsymbol{x}_\ell) = \{f(C)\boldsymbol{x}_\ell, f \in \mathbb{P}_{m-1}\}$, and

$$\begin{aligned}\rho_{\ell+1} &= \min_{x\in\mathcal{K}_m} \frac{x^{\mathsf{H}}Ax}{x^{\mathsf{H}}Bx} = \rho_\ell + \min_{x\in\mathcal{K}_m} \frac{x^{\mathsf{H}}(A-\rho_\ell B)x}{x^{\mathsf{H}}Bx} \\ &= \rho_\ell + \min_{f\in\mathbb{P}_{m-1}} \frac{\boldsymbol{x}_\ell^{\mathsf{H}} f(C) C f(C) \boldsymbol{x}_\ell}{\boldsymbol{x}_\ell^{\mathsf{H}} f(C) B f(C) \boldsymbol{x}_\ell}. \end{aligned} \tag{6}$$

3) Let $f_{\mathrm{opt}} \in \mathbb{P}_{m-1}$ be the minimizing polynomial that defines ϵ_m. Then $f_{\mathrm{opt}}(\omega_1) = 1$ by the definition, and also $\epsilon_m = \max_{j>1} |f_{\mathrm{opt}}(\omega_j)| < 1$ because

$$f(t) = \frac{t - (\omega_2+\omega_n)/2}{\omega_1 - (\omega_2+\omega_n)/2} \in \mathbb{P}_{m-1},\ f(\omega_1) = 1,\ \max_{j>1} |f(\omega_j)| = |f(\omega_2)| < 1.$$

4) $\boldsymbol{x}_\ell^{\mathsf{H}}(A - \rho_\ell B)\boldsymbol{x}_\ell = 0 \Rightarrow$ that $v_1^{\mathsf{H}}\boldsymbol{x}_\ell \neq 0$ and hence $f_{\mathrm{opt}}(C)\boldsymbol{x}_\ell \neq 0$ (why?). Thus

$$\begin{aligned}\rho_{\ell+1} &\le \rho_\ell + \frac{\boldsymbol{x}_\ell^{\mathsf{H}} f_{\mathrm{opt}}(C) C f_{\mathrm{opt}}(C) \boldsymbol{x}_\ell}{\boldsymbol{x}_\ell^{\mathsf{H}} f_{\mathrm{opt}}(C) B f_{\mathrm{opt}}(C) \boldsymbol{x}_\ell} = \rho_\ell + \frac{\boldsymbol{x}_\ell^{\mathsf{H}} V f_{\mathrm{opt}}^2(\Omega)\Omega V^{\mathsf{H}} \boldsymbol{x}_\ell}{\boldsymbol{x}_\ell^{\mathsf{H}} V f_{\mathrm{opt}}(\Omega) B_1 f_{\mathrm{opt}}(\Omega) V^{\mathsf{H}} \boldsymbol{x}_\ell} \\ &= \rho_\ell + \frac{y^{\mathsf{H}} f_{\mathrm{opt}}^2(\Omega)\Omega y}{y^{\mathsf{H}} f_{\mathrm{opt}}(\Omega) B_1 f_{\mathrm{opt}}(\Omega) y}, \qquad \text{where } B_1 = V^{\mathsf{H}} B V,\ y = V^{\mathsf{H}}\boldsymbol{x}_\ell. \end{aligned}$$

Proof

5) Write $y = [\xi_1, \xi_2, \ldots, \xi_n]^{\mathsf{T}} = \xi_1 e_1 + \hat{y}$, $\hat{y} = [0, \xi_2, \ldots, \xi_n]^{\mathsf{T}}$. We have

$$\begin{aligned} y^{\mathsf{H}} f_{\text{opt}}(\Omega) B_1 f_{\text{opt}}(\Omega) y &= (\xi_1 e_1 + \hat{y})^{\mathsf{H}} f_{\text{opt}}(\Omega) B_1 f_{\text{opt}}(\Omega)(\xi_1 e_1 + \hat{y}) \\ &= \xi_1^2 f_{\text{opt}}(\omega_1)^2 e_1^{\mathsf{H}} B_1 e_1 + 2\xi_1 f_{\text{opt}}(\omega_1) e_1^{\mathsf{H}} B_1 f_{\text{opt}}(\Omega)\hat{y} \\ &\quad + \hat{y}^{\mathsf{H}} f_{\text{opt}}(\Omega) B_1 f_{\text{opt}}(\Omega)\hat{y} \\ &= \xi_1^2 \beta_1^2 + 2\xi_1 \beta_2 + \beta_3^2, \end{aligned}$$

where

$$\begin{aligned} \beta_1^2 &= e_1^{\mathsf{H}} B_1 e_1 = v_1^{\mathsf{H}} B v_1, \\ \beta_3^2 &= \hat{y}^{\mathsf{H}} f_{\text{opt}}(\Omega) B_1 f_{\text{opt}}(\Omega)\hat{y} \le \max_{j>1} f_{\text{opt}}(\omega_j)^2 \|B_1\|_2 \|\hat{y}\|_2^2 = \epsilon_m^2 \|B\|_2 \|\hat{y}\|_2^2, \\ |\beta_2| &= |e_1^{\mathsf{H}} B_1 f_{\text{opt}}(\Omega)\hat{y}| \le \beta_1 \beta_3. \end{aligned}$$

Note $\sum_j \omega_j \xi_j^2 = y^{\mathsf{H}} \Omega y = \boldsymbol{x}_\ell^{\mathsf{H}} (A - \boldsymbol{\rho}_\ell B)\boldsymbol{x}_\ell = 0 \Rightarrow |\omega_1|\xi_1^2 = \sum_{j>1} \omega_j \xi_j^2 \ge \omega_2 \|\hat{y}\|_2^2$.
Hence

$$\beta_3 \le \epsilon_m \|B\|_2^{1/2} \left(\frac{|\omega_1|}{\omega_2} \right)^{1/2} |\xi_1|.$$

Proof

6) $y^{\mathsf{H}} f_{\text{opt}}^2(\Omega)\Omega y = \xi_1^2\omega_1 + \hat{y}^{\mathsf{H}} f_{\text{opt}}(\Omega)^2\Omega\hat{y}$, and

$$y^{\mathsf{H}} f_{\text{opt}}^2(\Omega)\Omega y = \sum_j \omega_j f_{\text{opt}}^2(\omega_j)\xi_\ell^2 \le \sum_j \omega_j \xi_j^2 = y^{\mathsf{H}}\Omega y = 0,$$

$$\hat{y}^{\mathsf{H}} f_{\text{opt}}^2(\Omega)\Omega\hat{y} = \sum_{j>1} \omega_j f_{\text{opt}}^2(\omega_j)\xi_\ell^2 \le \epsilon_m^2 \sum_{j>1} \omega_j \xi_j^2 = \epsilon_m^2|\omega_1|\xi_1^2.$$

$$\begin{aligned}
\frac{y^{\mathsf{H}} f_{\text{opt}}^2(\Omega)\Omega y}{y^{\mathsf{H}} f_{\text{opt}}(\Omega) B_1 f_{\text{opt}}(\Omega) y} &\le \frac{\xi_1^2\omega_1 + \hat{y}^{\mathsf{H}} f_{\text{opt}}(\Omega)^2\Omega\hat{y}}{\xi_1^2\beta_1^2 + 2|\xi_1|\beta_1\beta_3 + \beta_3^2} \\
&= \frac{\omega_1}{\beta_1^2} - \frac{\omega_1}{\beta_1^2}\cdot\frac{2|\xi_1|\beta_1\beta_3 + \beta_3^2}{\xi_1^2\beta_1^2 + 2|\xi_1|\beta_1\beta_3 + \beta_3^2} + \frac{\hat{y}^{\mathsf{H}} f_{\text{opt}}(\Omega)^2\Omega\hat{y}}{\xi_1^2\beta_1^2 + 2|\xi_1|\beta_1\beta_3 + \beta_3^2} \\
&\le \frac{\omega_1}{\beta_1^2} - \frac{\omega_1}{\beta_1^2}\cdot\frac{2|\xi_1|\beta_1\beta_3}{\xi_1^2\beta_1^2} + \frac{\hat{y}^{\mathsf{H}} f_{\text{opt}}(\Omega)^2\Omega\hat{y}}{\xi_1^2\beta_1^2} \\
&\le \frac{\omega_1}{\beta_1^2} + 2\left(\frac{|\omega_1|}{\beta_1^2}\right)^{3/2}\epsilon_m\left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + \frac{|\omega_1|}{\beta_1^2}\epsilon_m^2.
\end{aligned}$$

Proof

7) We have proved

$$\frac{y^{\mathsf{H}} f_{\text{opt}}^2(\Omega)\Omega y}{y^{\mathsf{H}} f_{\text{opt}}(\Omega) B_1 f_{\text{opt}}(\Omega) y} \leq \frac{\omega_1}{\beta_1^2} + 2\left(\frac{|\omega_1|}{\beta_1^2}\right)^{3/2} \epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + \frac{|\omega_1|}{\beta_1^2}\epsilon_m^2.$$

Note $\frac{\omega_1}{\beta_1^2} = \omega_1 \frac{v_1^{\mathsf{H}} v_1}{v_1^{\mathsf{H}} B v_1} = (\lambda_1 - \boldsymbol{\rho}_\ell) + \mathcal{O}(|\lambda_1 - \boldsymbol{\rho}_\ell|^2)$ by the lemma. Therefore

$$\begin{aligned}
\boldsymbol{\rho}_{\ell+1} - \lambda_1 &\leq \boldsymbol{\rho}_\ell - \lambda_1 + \frac{y^{\mathsf{H}} f_{\text{opt}}^2(\Omega)\Omega y}{y^{\mathsf{H}} f_{\text{opt}}(\Omega) B_1 f_{\text{opt}}(\Omega) y} \\
&\leq \underbrace{\boldsymbol{\rho}_\ell - \lambda_1 + \frac{\omega_1}{\beta_1^2}}_{\delta_\ell} + 2\left(\frac{|\omega_1|}{\beta_1^2}\right)^{3/2} \epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + \frac{|\omega_1|}{\beta_1^2}\epsilon_m^2 \\
&\leq \delta_\ell + 2(\boldsymbol{\rho}_\ell - \lambda_1)^{3/2} \epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + (\boldsymbol{\rho}_\ell - \lambda_1)\epsilon_m^2,
\end{aligned}$$

as expected. □

What is Preconditioning?

Preconditioning. Transforming a problem that is "easier" (e.g., taking less time) to solve iteratively.

Preconditioning natural for linear systems: transform $Ax = b$ to $KAx = Kb$ which is "easier" than before. Extreme case: $KA = I$, i.e., $K = A^{-1}$, then $x = Kb$. But this is impractical!

A comprise: make $KA \approx I$ as much as practical. Here $KA \approx I$ is understood either $\|KA - I\|$ is relatively small or $KA - I$ is near a low rank matrix.

Preconditioning not so natural for eigenvalue problems: transform $A - \lambda B$ to $KA - \lambda KB$ or $L^{\mathrm{H}}AL - \lambda L^{\mathrm{H}}BL$ which is "easier" than before.

- No straightforward explanation as to what K makes $KA - \lambda KB$ "easier"
- No straightforward explanation as to what L makes $L^{\mathrm{H}}AL - \lambda L^{\mathrm{H}}BL$ "easier", except L being the eigenvector matrix that is unknown. No easy way to approximate the unknown eigenvector matrix either.

Will present two ways to understand eigen-problem preconditioning and construct preconditioners.

Eigen-problem preconditioning, I

Ideal search direction p: starting at $\boldsymbol{x}_\ell$, p points to the optimum, i.e., the optimum is on the line $\{\boldsymbol{x}_\ell + tp \,:\, t \in \mathbb{C}\}$. How can it be done with unknown optimum?

Expand $\boldsymbol{x}_\ell$ as a linear combination of u_ℓ

$$\boldsymbol{x}_\ell = \sum_{j=1}^{n} \alpha_j u_j =: \alpha_1 u_1 + \boldsymbol{v}, \quad \boldsymbol{v} = \sum_{j=2}^{n} \alpha_j u_j \perp_B u_1.$$

Then ideal $p = \alpha u_1 + \beta \boldsymbol{v}$, $\beta \neq 0$ such that $\alpha_1 \beta - \alpha \neq 0$ (otherwise $p = \beta \boldsymbol{x}_\ell$).

Ideal p has to be approximated to be practical. One such approximate p is

$$p = (A - \sigma B)^{-1} \boldsymbol{r}_\ell = (A - \sigma B)^{-1} [A - \boldsymbol{\rho}_\ell B] \boldsymbol{x}_\ell,$$

where $\boldsymbol{\rho}_\ell \neq \sigma \approx \lambda_1$, also reasonably we assume $\sigma \neq \lambda_j$ for all $j > 1$. Why so?

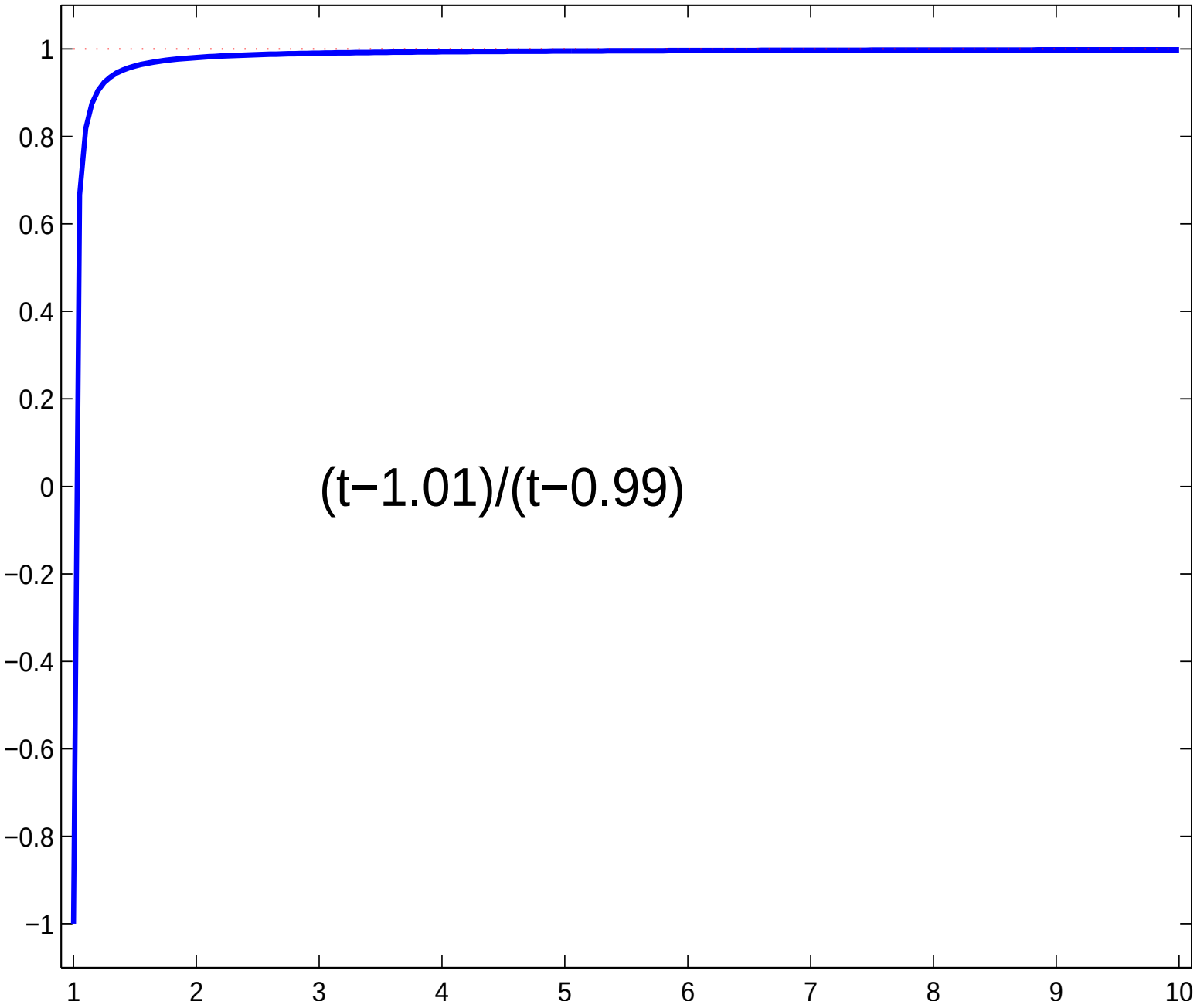
$$p = \sum_{j=1}^{n} \mu_j \alpha_j u_j, \quad \mu_j := \frac{\lambda_j - \boldsymbol{\rho}_\ell}{\lambda_j - \sigma}.$$

Now if $\lambda_1 \leq \boldsymbol{\rho}_\ell < \lambda_2$ and if the gap $\lambda_2 - \lambda_1$ is reasonably modest, then

$$\mu_j \approx 1 \quad \text{for } j > 1$$

to give a $p \approx \alpha v_1 + \boldsymbol{v}$, resulting in fast convergence.

$\mu_j = (\lambda_j - \boldsymbol{\rho}_\ell)/(\lambda_j - \sigma)$



Eigen-problem preconditioning, I

Preconditioner $(A - \sigma B)^{-1}$

Let $\boldsymbol{x}_\ell = \sum_{j=1}^{n} \alpha_j u_j$, and suppose $\alpha_1 \neq 0$. If $\sigma \neq \boldsymbol{\rho}_\ell$ such that

$$\text{either } \mu_1 < \mu_j \text{ for } 2 \leq j \leq n \text{ or } \mu_1 > \mu_j \text{ for } 2 \leq j \leq n,$$

where $\mu_j = \frac{\lambda_j - \boldsymbol{\rho}_\ell}{\lambda_j - \sigma}$, then

$$\tan\theta_B(u_1, \mathcal{K}_m) \leq 2\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-1} \tan\theta_B(u_1, \boldsymbol{x}_\ell),$$
$$0 \leq \boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq 4\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-2} \tan\theta_B(u_1, \boldsymbol{x}_\ell),$$

where $\mathcal{K}_m := \mathcal{K}_m([A - \sigma B]^{-1}[A - \boldsymbol{\rho}_\ell B], \boldsymbol{x}_\ell)$, and

$$\eta = \begin{cases} \dfrac{\lambda_n - \sigma}{\lambda_n - \lambda_1} \cdot \dfrac{\lambda_2 - \lambda_1}{\lambda_2 - \sigma}, & \text{if } \mu_1 < \mu_j \text{ for } 2 \leq j \leq n, \\ \dfrac{\lambda_2 - \sigma}{\lambda_2 - \lambda_1} \cdot \dfrac{\lambda_n - \lambda_1}{\lambda_n - \sigma}, & \text{if } \mu_1 > \mu_j \text{ for } 2 \leq j \leq n. \end{cases}$$

Discussion

$\eta \approx 1$ (fast convergence) if $\sigma \approx \lambda_1$. In fact $\eta = 1$ (implying $\Delta_\eta = \infty$) if $\sigma = \lambda_1$.

But shift σ needs to make μ_1 either smallest or biggest among all μ_j. Three interesting cases:

- $\sigma < \lambda_1 \leq \rho < \lambda_2$, μ_1 smallest
- $\lambda_1 < \sigma < \rho < \lambda_2$, μ_1 biggest
- $\lambda_1 < \rho < \sigma < \lambda_2$, μ_1 smallest.

$(A - \sigma B)^{-1}$ realized through linear system solving, but cost is high if solved accurately, thus only approximately, such as

- incomplete decompositions LDL^{H} of $A - \sigma B$ with/without an iterative method
- CG, MINRES

(more from Sherry Li's lectures.)

Eigen-problem preconditioning, II

Use Golub and Ye's Theorem (2002) as starting point:

$$\boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq (\boldsymbol{\rho}_\ell - \lambda_1)\epsilon_m^2 + 2(\boldsymbol{\rho}_\ell - \lambda_1)^{3/2}\epsilon_m \left(\frac{\|B\|_2}{\omega_2}\right)^{1/2} + \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2),$$

$$\epsilon_m := \min_{f \in \mathbb{P}_{m-1}, f(\omega_1)=1} \max_{j>1} |f(\omega_j)|,$$

where $\omega_1 < \omega_2 \leq \cdots \leq \omega_n$ are the eigenvalues of $A - \boldsymbol{\rho}_\ell B$.

Idea: Transform $A - \lambda B$ to $L^{-1}(A - \lambda B)L^{-\mathsf{H}}$ so that $L^{-1}(A - \boldsymbol{\rho}_\ell B)L^{-\mathsf{H}}$ has "better" eigenvalue distribution, i.e., much smaller ϵ_m.

Ideal: $\omega_2 = \cdots = \omega_n$, then $\epsilon_m = 0$ for $m \geq 2$ and thus $\boldsymbol{\rho}_{\ell+1} - \lambda_1 = \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2)$, quadratic convergence.

$A - \boldsymbol{\rho}_\ell B = LDL^{\mathsf{H}} \Rightarrow L^{-1}(A - \boldsymbol{\rho}_\ell B)L^{-\mathsf{H}} = D = \mathrm{diag}(\pm 1)$. Ideal but not practical:

1. $A - \boldsymbol{\rho}_\ell B = LDL^{\mathsf{H}}$ may not exist at all. It exists if all leading principle submatrices are nonsingular.
2. $A - \boldsymbol{\rho}_\ell B = LDL^{\mathsf{H}}$ may not be numerically stable to compute, especially when $\boldsymbol{\rho}_\ell \approx \lambda_1$.
3. L significantly denser than A and B combined. Ensuing computations are too expensive.

Eigen-problem preconditioning, II

Compromise. $A - \boldsymbol{\rho}_\ell B \approx LDL^{\mathsf{H}}$ with a good chance that

> one smallest isolated eigenvalue ω_1, and the rest ω_j $(2 \le j \le n)$
> a few tight clusters.

Here $A - \boldsymbol{\rho}_\ell B \approx LDL^{\mathsf{H}}$ includes not only the usual "approximately equal", but also when $(A - \boldsymbol{\rho}_\ell B) - LDL^{\mathsf{H}}$ approximately of a low rank.

L varies from one iterative step to another; Can be expensive; Possible to use constant preconditioner, i.e., one L for all steps or change it every few steps.

Constant preconditioner: Use a shift $\sigma \approx \lambda_1$, and perform an incomplete LDL^{H} decomposition of $A - \sigma B \approx LDL^{\mathsf{H}}$. Then

$$\widehat{C}_\ell = L^{-1}(A - \sigma B)L^{-\mathsf{H}} + (\sigma - \boldsymbol{\rho}_\ell)L^{-1}BL^{-\mathsf{H}} \approx D$$

would have a better spectral distribution so long as $(\sigma - \boldsymbol{\rho}_\ell)L^{-1}BL^{-\mathsf{H}}$ is small relative to $\widehat{C}_\ell$.

Insisted so far about applying ESD straightforwardly to the transformed problem $L^{-1}(A - \lambda B)L^{-\mathsf{H}}$. But there is alternative, perhaps better, way.

Eigen-problem preconditioning, II

$A - \lambda B$ to $\widehat{A}_\ell - \lambda \widehat{B}_\ell := L_\ell^{-1}(A - \lambda B)L_\ell^{-\mathsf{H}}$. Typical step of ESD for $\widehat{A}_\ell - \lambda \widehat{B}_\ell$:

> compute the smallest eigenvalue μ and corresponding eigenvector v of $\widehat{Z}^{\mathsf{H}}(\widehat{A}_\ell - \lambda \widehat{B}_\ell)\widehat{Z}$, where $\widehat{Z} \in \mathbb{C}^{n\times m}$ is a basis matrix of Krylov subspace $\mathcal{K}_m(\widehat{A}_\ell - \hat{\boldsymbol{\rho}}_\ell \widehat{B}_\ell, \hat{\boldsymbol{x}}_\ell)$.

Notice $\left[\widehat{A}_\ell - \hat{\boldsymbol{\rho}}_\ell \widehat{B}_\ell\right]^j \hat{\boldsymbol{x}}_\ell = L_\ell^{\mathsf{H}} \left[(L_\ell L_\ell^{\mathsf{H}})^{-1}(A - \hat{\boldsymbol{\rho}}_\ell B)\right]^j (L_\ell^{-\mathsf{H}} \hat{\boldsymbol{x}}_\ell)$ to see

$$L_\ell^{-\mathsf{H}} \cdot \mathcal{K}_m(\widehat{A}_\ell - \hat{\boldsymbol{\rho}}_\ell \widehat{B}_\ell, \hat{\boldsymbol{x}}_\ell) = \mathcal{K}_m(\, K_\ell(A - \hat{\boldsymbol{\rho}}_\ell B), \boldsymbol{x}_\ell), \;\; \boldsymbol{x}_\ell = L_\ell^{-\mathsf{H}} \hat{\boldsymbol{x}}_\ell, \;\; K_\ell = (L_\ell L_\ell^{\mathsf{H}})^{-1}.$$

So $Z = L_\ell^{-\mathsf{H}} \widehat{Z}$ is a basis matrix of Krylov subspace $\mathcal{K}_m(\, K_\ell(A - \hat{\boldsymbol{\rho}}_\ell B), \boldsymbol{x}_\ell)$. Also

$$\widehat{Z}^{\mathsf{H}}(\widehat{A}_\ell - \lambda \widehat{B}_\ell)\widehat{Z} = (L_\ell^{-\mathsf{H}} \widehat{Z})^{\mathsf{H}}(A - \lambda B)(L_\ell^{-\mathsf{H}} \widehat{Z}) = Z^{\mathsf{H}}(A - \lambda B)Z,$$

$$\hat{\boldsymbol{\rho}}_\ell = \frac{\hat{\boldsymbol{x}}_\ell^{\mathsf{H}} \widehat{A}_\ell \hat{\boldsymbol{x}}_\ell}{\hat{\boldsymbol{x}}_\ell^{\mathsf{H}} \widehat{B}_\ell \hat{\boldsymbol{x}}_\ell} = \frac{\boldsymbol{x}_\ell^{\mathsf{H}} A \boldsymbol{x}_\ell}{\boldsymbol{x}_\ell^{\mathsf{H}} B \boldsymbol{x}_\ell} = \boldsymbol{\rho}_\ell.$$

The typical step can be reformulated equivalently to

> compute the smallest eigenvalue μ and corresponding eigenvector v of $Z^{\mathsf{H}}(A - \lambda B)Z$, where $Z \in \mathbb{C}^{n\times m}$ is a basis matrix of Krylov subspace $\mathcal{K}_m(\, K_\ell(A - \boldsymbol{\rho}_\ell B), \boldsymbol{x}_\ell)$, where $K_\ell = (L_\ell L_\ell^{\mathsf{H}})^{-1}$.

Extended Preconditioned SD

Extended Preconditioned Steepest Descent method

Given an initial approximation $\boldsymbol{x}_0$ to u_1, a relative tolerance `rtol`, and an integer $m \geq 2$, the algorithm attempts to compute an approximate eigenpair to (λ_1, u_1) with the prescribed `rtol`.

$\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|_B$, $\boldsymbol{\rho}_0 = \boldsymbol{x}_0^{\mathsf{H}} A\boldsymbol{x}_0$, $\boldsymbol{r}_0 = A\boldsymbol{x}_0 - \boldsymbol{\rho}_0 B\boldsymbol{x}_0$;
for $\ell = 0, 1, \ldots$ **do**
  if $\|\boldsymbol{r}_\ell\|_2/(\|A\boldsymbol{x}_\ell\|_2 + |\boldsymbol{\rho}_\ell|\,\|B\boldsymbol{x}_\ell\|_2) \leq$ `rtol` **then**
    BREAK;
  else
    construct a preconditioner K_ℓ;
    compute a basis matrix $Z \in \mathbb{C}^{n\times m}$ of Krylov subspace $\mathcal{K}_m(\, K_\ell(A - \boldsymbol{\rho}_\ell B), \boldsymbol{x}_\ell)$;
    compute the smallest eigenvalue μ and corresponding eigenvector v of $Z^{\mathsf{H}}(A - \lambda B)Z$;
    $y = Zv$, $\boldsymbol{x}_{\ell+1} = y/\|y\|_B$;
    $\boldsymbol{\rho}_{\ell+1} = \mu$, $\boldsymbol{r}_{\ell+1} = A\boldsymbol{x}_{\ell+1} - \boldsymbol{\rho}_{\ell+1} B\boldsymbol{x}_{\ell+1}$;
  end if
end for
return $(\boldsymbol{\rho}_\ell, \boldsymbol{x}_\ell)$ as an approximate eigenpair to (λ_1, u_1).

It actually includes Eigen-problem preconditioning, I & II.

Convergence Rate

Theorem on Convergence Rate (Golub & Ye, 2002)

Suppose λ_1 is simple, i.e., $\lambda_1 < \lambda_2$, and $\lambda_1 < \boldsymbol{\rho}_\ell < \lambda_2$, and preconditioner $K_\ell \succ 0$. Let

$$\omega_1 < \omega_2 \leq \cdots \leq \omega_n$$

be the eigenvalues of $K_\ell(A - \boldsymbol{\rho}_\ell B)$ and v_1 be an eigenvector corresponding to ω_1. Then

$$\boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq (\boldsymbol{\rho}_\ell - \lambda_1)\epsilon_m^2 + 2(\boldsymbol{\rho}_\ell - \lambda_1)^{3/2}\epsilon_m \left(\frac{\|L_\ell^{-1} B L_\ell^{-\mathsf{H}}\|_2}{\omega_2} \right)^{1/2} + \delta_\ell$$

where

$$0 \leq \delta_\ell := \boldsymbol{\rho}_\ell - \lambda_1 + \omega_1 \frac{v_1^{\mathsf{H}} K_\ell^{-1} v_1}{v_1^{\mathsf{H}} B v_1} = \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2),$$

$$\epsilon_m := \min_{f \in \mathbb{P}_{m-1}, f(\omega_1)=1} \max_{j>1} |f(\omega_j)|.$$

Discussion

$\epsilon_m := \min_{f\in\mathbb{P}_{m-1}, f(\omega_1)=1} \max_{j>1} |f(\omega_j)|$ usually unknown, except,

1. $m = 2$ for which the optimal f_{opt} is

$$f(t) = \frac{t - (\omega_2+\omega_n)/2}{\omega_1 - (\omega_2+\omega_n)/2} \in \mathbb{P}_{m-1},\ f(\omega_1) = 1,\ \max_{j>1} |f(\omega_j)| = |f(\omega_2)| < 1.,$$

$$\epsilon_2 = \frac{1-\eta}{1+\eta}, \quad \eta = \frac{\omega_2 - \omega_1}{\omega_n - \omega_1}.$$

2. $\omega_2 = \cdots = \omega_n$ for which $f_{\text{opt}}(t) = (t-\omega_2)/(\omega_1 - \omega_2)$ and $\epsilon_m = 0$ for all $m \geq 2$.

In general ϵ_m can be bounded by using the Chebyshev polynomial

$$f(t) = \mathscr{T}_{m-1}\left(\frac{2t - (\omega_n+\omega_2)}{\omega_n - \omega_2}\right) \Big/ \mathscr{T}_{m-1}\left(\frac{1+\eta}{1-\eta}\right), \qquad f(\omega_1) = 1,$$

$$\epsilon_m \leq \max_{\omega_2 \leq t \omega_n} |f(t)| = \left[\mathscr{T}_{m-1}\left(\frac{1+\eta}{1-\eta}\right)\right]^{-1} = 2\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-1}.$$

Discussion (cont'd)

$$\boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq (\boldsymbol{\rho}_\ell - \lambda_1)\epsilon_m^2 + 2(\boldsymbol{\rho}_\ell - \lambda_1)^{3/2}\epsilon_m \left(\frac{\|L_\ell^{-1} B L_\ell^{-\mathsf{H}}\|_2}{\omega_2} \right)^{1/2} + \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2).$$

Ignoring high order terms, $\dfrac{\boldsymbol{\rho}_{\ell+1} - \lambda_1}{\boldsymbol{\rho}_\ell - \lambda_1} \lessapprox \epsilon_m^2$.

If $\epsilon_m = 0$ (unlikely, however), then $\boldsymbol{\rho}_{\ell+1} - \lambda_1 = \mathcal{O}(|\boldsymbol{\rho}_\ell - \lambda_1|^2)$, quadratically convergence.

Locally, $\boldsymbol{\rho}_\ell \approx \lambda_1$, $\mathrm{eig}(K_\ell(A - \boldsymbol{\rho}_\ell B)) \approx \mathrm{eig}(K_\ell(A - \lambda_1 B)) = \{0 = \gamma_1 < \gamma_2 \leq \cdots \leq \gamma_n\}$, and

$$\epsilon_m \approx \min_{f \in \mathbb{P}_{m-1}, f(\gamma_1)=1} \max_{j>1} |f(\gamma_j)| \leq 2\left[\Delta_\eta^{m-1} + \Delta_\eta^{-(m-1)}\right]^{-1},$$

$$\Delta_\eta = \frac{1+\sqrt{\eta}}{1-\sqrt{\eta}}, \quad \eta = \frac{\gamma_2 - \gamma_1}{\gamma_n - \gamma_1}.$$

Convergence Rate: Samokish Theorem

Theorem on Convergence Rate (Samokish, 1958)

$m = 2$, $\text{eig}(K_\ell(A - \lambda_1 B)) = \{0 = \gamma_1 < \gamma_2 \leq \cdots \leq \gamma_n\}$. Suppose λ_1 is simple, i.e., $\lambda_1 < \lambda_2$, and $\lambda_1 < \boldsymbol{\rho}_\ell < \lambda_2$,

$$\delta = \sqrt{\|B^{1/2} K_\ell B^{1/2}\|_2 \, [\boldsymbol{\rho}_\ell - \lambda_1]}, \quad \tau = \frac{2}{\gamma_2 + \gamma_n}.$$

If $\tau \left(\sqrt{\gamma_n} + \delta\right) \delta < 1$, then

$$\boldsymbol{\rho}_{\ell+1} - \lambda_1 \leq \left[\frac{\epsilon_2 + \tau\sqrt{\gamma_n}\,\delta}{1 - \tau(\sqrt{\gamma_n} + \delta)\delta}\right]^2 [\boldsymbol{\rho}_\ell - \lambda_1], \tag{7}$$

where

$$\eta = \frac{\gamma_2 - \gamma_1}{\gamma_n - \gamma_1}, \quad \epsilon_2 = 2\left[\Delta_\eta + \Delta_\eta^{-1}\right]^{-1} = \frac{1 - \eta}{1 + \eta}.$$

Asymptotically $\boldsymbol{\rho}_{\ell+1} - \lambda_1 \lessapprox \epsilon_2^2 \, (\boldsymbol{\rho}_\ell - \lambda_1)$, same as Golub & Ye (2002), but (7) is strict.

Proof

1) $t_{\text{opt}} = \arg\min_t \rho(\boldsymbol{x}_\ell + tK_\ell \boldsymbol{r}_\ell)$, $y = \boldsymbol{x}_\ell + t_{\text{opt}} K_\ell \boldsymbol{r}_\ell$. Thus $\boldsymbol{\rho}_{\ell+1} = \rho(y)$.
2) Drop subscript ℓ to $\boldsymbol{x}$, $\boldsymbol{r}$, and K: $\boldsymbol{r}_\ell = r(\boldsymbol{x})$, $\boldsymbol{\rho}_\ell = \rho(\boldsymbol{x})$.
3) $z = \boldsymbol{x} - \tau K r(\boldsymbol{x})$. Then $\lambda_1 \le \rho(y) \le \rho(z)$, thus $\rho(y) - \lambda_1 \le \rho(z) - \lambda_1$.
4) Suffices to show $\rho(z) - \lambda_1 \le$ RHS of (7).
5) $A - \lambda_1 B$ is symmetric positive semidefinite. $\|\cdot\|_{A-\lambda_1 B}$ is a semi-norm.

$$\|w\|^2_{A-\lambda_1 B} = [\rho(w) - \lambda_1]\|w\|^2_B,$$
$$\|[I - \tau K(A - \lambda_1 B)]w\|_{A-\lambda_1 B} \le \epsilon_2 \|w\|_{A-\lambda_1 B}.$$

6) Write $z = [I - \tau K(A - \lambda_1 B)]\boldsymbol{x} + \tau[\rho(\boldsymbol{x}) - \lambda_1]KB\boldsymbol{x}$, and assume $\|\boldsymbol{x}\|_B = 1$.

$$\begin{aligned}
\|z\|_{A-\lambda_1 B} &= \sqrt{\rho(z) - \lambda_1}\|z\|_B, \\
\|z\|_{A-\lambda_1 B} &\le \|[I - \tau K(A - \lambda_1 B)]\boldsymbol{x}\|_{A-\lambda_1 B} + \tau[\rho(\boldsymbol{x}) - \lambda_1]\|KB\boldsymbol{x}\|_{A-\lambda_1 B} \\
&\le \epsilon_2 \|\boldsymbol{x}\|_{A-\lambda_1 B} + \tau[\rho(\boldsymbol{x}) - \lambda_1]\sqrt{\gamma_n}\|B\boldsymbol{x}\|_K \\
&\le \epsilon_2 \sqrt{\rho(\boldsymbol{x}) - \lambda_1} + \tau[\rho(\boldsymbol{x}) - \lambda_1]\sqrt{\gamma_n \|B^{1/2} K B^{1/2}\|_2} \\
&= (\epsilon_2 + \tau\sqrt{\gamma_n}\,\delta)\sqrt{\rho(\boldsymbol{x}) - \lambda_1}.
\end{aligned}$$

7) $z = \boldsymbol{x} - \tau K r(\boldsymbol{x})$, and $\|\boldsymbol{x}\|_B = 1$.

$$
\begin{aligned}
\|z\|_B &\geq \|\boldsymbol{x}\|_B - \tau \|Kr(\boldsymbol{x})\|_B = 1 - \tau \|Kr(\boldsymbol{x})\|_B, \\
\|Kr(\boldsymbol{x})\|_B &= \|K(A - \lambda_1 B)\boldsymbol{x} - [\rho(\boldsymbol{x}) - \lambda_1] KB\boldsymbol{x}\|_B \\
&\leq \|K(A - \lambda_1 B)\boldsymbol{x}\|_B + [\rho(\boldsymbol{x}) - \lambda_1] \|KB\boldsymbol{x}\|_B \\
&\leq \sqrt{\|K^{1/2} B K^{1/2}\|_2 \gamma_n}\, \|\boldsymbol{x}\|_{A-\lambda_1 B} + [\rho(\boldsymbol{x}) - \lambda_1] \|B^{1/2} K B^{1/2}\|_2 \|\boldsymbol{x}\|_B \\
&= \sqrt{\gamma_n}\delta + \delta^2.
\end{aligned}
$$

8) Finally use

$$
\rho(z) - \lambda_1 = \frac{\|z\|^2_{A-\lambda_1 B}}{\|z\|^2_B} \leq \frac{\|z\|^2_{A-\lambda_1 B}}{[1 - \tau \|Kr(\boldsymbol{x})\|_B]^2}
$$

to complete the proof.

Deflation

So far computing (λ_1, u_1) by *single-vector* steepest descent type methods.

To compute any following eigenpairs, must incorporate deflation techniques.

Or use a multi-vector/block method (not discussed yet). Deflation is also a necessary tool to make a block method more efficient.

Assume acceptable approximations to (λ_i, u_i) for $1 \leq j \leq k$ known.
Diagonal $\boldsymbol{D} \in \mathbb{R}^{k \times k}$ holds known approximations of λ_i,
$\boldsymbol{U} \in \mathbb{R}^{n \times k}$ holds known approximations of u_i.
Assume $\boldsymbol{U}^{\mathsf{H}} B \boldsymbol{U} = I_k$.

Deflation: avoid computing (λ_i, u_i) for $1 \leq j \leq k$, and seek approximation to (λ_{k+1}, u_{k+1}). Will discuss two deflation techniques.

Through orthogonalizing against $\boldsymbol{U}$

When the basis matrix Z is computed, make sure that Z is B-orthogonal to $\boldsymbol{U}$. E.g., build a basis matrix Z for $\mathcal{K}_m(K(A-\rho B), x)$ such that $\boldsymbol{U}^{\mathrm{H}} \perp_B Z = 0$. Suppose $x \perp_B \boldsymbol{U} = 0$ already.

Arnoldi-like process

1: $Z_{(:,1)} = x/\|x\|_B$, $\rho = x^{\mathrm{H}}Ax/\|x\|_B^2$;
2: **for** $i = 2$ to m **do**
3: $\quad q = K(AZ_{(:,i-1)} - \rho BZ_{(:,i-1)})$;
4: $\quad q = q - \boldsymbol{U}(\boldsymbol{U}^{\mathrm{H}}(Bq))$;
5: $\quad$ **for** $j = 1$ to $i-1$ **do**
6: $\quad\quad t = Z_{(:,i-1)}^{\mathrm{H}} Bq$, $q = q - Z_{(:,i-1)}t$;
7: $\quad$ **end for**
8: $\quad t = \|q\|_B$;
9: $\quad$ **if** $t > 0$ **then**
10: $\quad\quad Z_{(:,i)} = q/t$;
11: $\quad$ **else**
12: $\quad\quad$ **BREAK**;
13: $\quad$ **end if**
14: **end for**

Note: Keep AZ for later use.

Through shifting λ_i away

Lemma

Let $U_1 = U_{(:,1:k)} = [u_1, \ldots, u_k]$. $(A + \zeta B U_1 U_1^{\mathrm{H}} B) - \lambda B$ and $A - \lambda B$ share same eigenvectors u_i, but the eigenvalues of $(A + \zeta B U_1 U_1^{\mathrm{H}} B) - \lambda B$ are

$$\lambda_i + \zeta \text{ for } 1 \leq i \leq k \text{ and } \lambda_i \text{ for } k+1 \leq i \leq n.$$

Modify $A - \lambda B$ in form, but not explicitly, to $(A + \zeta B \boldsymbol{U}\boldsymbol{U}^{\mathrm{H}} B) - \lambda B$, where ζ should be selected such that $\zeta + \lambda_1 \geq \lambda_{k+2}$.

But λ_{k+2} is unknown. What we can do in practice to pick ζ a sufficiently large number.

Block Steepest Descent Method

- Block Steepest Descent Method
- Block Extended Steepest Descent Method
- Block Preconditioned Extended Steepest Descent Method

Why block versions

Single-vector SD and variations:

- compute (λ_1, u_1), and, with deflations, other (λ_i, u_i), one pair at a time. Most computations are of matrix-vector type.
- Slow convergence if $(\gamma_2 - \gamma_1)/(\gamma_n - \gamma_1)$ tiny; usually happens when λ_2 very close to λ_1. (γ_i are eigenvalues of $K(A - \lambda_1 B)$.)
- Often in practice, there are needs to compute the first few eigenpairs, not just the first one.

Block versions:

- Can simultaneously compute the first k eigenpairs (λ_j, u_j);
- Run more efficiently on modern computer architecture: more computations in matrix-matrix multiplication type;
- Better rates of convergence; can save overall cost by using a block size that is slightly bigger than the number of asked eigenpairs.

In summary, the benefits of using a block variation are similar to those of using the simultaneous subspace iteration *vs.* the power method.

Block SD

Start with $X_0 \in \mathbb{C}^{n\times n_b}$, $\operatorname{rank}(X_0) = n_b \geq k$, instead of just one vector $\boldsymbol{x}_0 \in \mathbb{C}^n$.

May assume jth column of X_0 approximates u_j; otherwise $\mathcal{R}(X_0)$ approximates $\operatorname{span}\{u_1, \ldots, u_{n_b}\}$. In the latter, preprocessing X_0:

1. compute eigen-decomposition $(X_0^{\mathsf{H}} A X_0)\, W = (X_0^{\mathsf{H}} B X_0)\, W\Omega$, where $\Omega = \operatorname{diag}(\rho_{0;1}, \rho_{0;2}, \ldots, \rho_{0;n_b})$;
2. Reset $X_0 := X_0 W$.

Can always assume jth column of X_0 approximates u_j.

Typical ℓth iterative step: already have

$$X_\ell = [x_{\ell;1}, x_{\ell;2}, \ldots, x_{\ell;n_b}] \in \mathbb{C}^{n\times n_b}, \quad j\text{th column } x_{\ell;j} \text{ approximates } u_j,$$
$$\Omega_\ell = \operatorname{diag}(\rho_{\ell;1}, \rho_{\ell;2}, \ldots, \rho_{\ell;n_b}), \qquad \rho_{\ell,j} = \rho(x_{\ell;j}) \approx \lambda_j.$$

To compute new approximations as follows.

1. Compute a basis matrix Z of $\mathcal{R}([X_\ell, R_\ell])$ by, e.g., MGS in the B-inner product, keeping in mind that X_ℓ is B-orthonormal already;
2. Find the first n_b eigenpairs of $Z^{\mathsf{H}} A Z - \lambda Z^{\mathsf{H}} B Z$ to get $(Z^{\mathsf{H}} A Z) W = (Z^{\mathsf{H}} B Z) W \Omega_{\ell+1}$, $\Omega_{\ell+1} = \operatorname{diag}(\rho_{\ell+1;1}, \rho_{\ell+1;2}, \ldots, \rho_{\ell+1;n_b})$;
3. Set $X_{\ell+1} = ZW$.

Block Extended SD

Block SD (previous slide) is is the stronger version of **Simultaneous Rayleigh Quotient Minimization Method** of Longsine and McCormick (1980).

Note that $r(x_{\ell;j}) = (A - \rho_{\ell;j}B)x_{\ell;j}$ and thus

$$\mathcal{R}([X_\ell, R_\ell]) = \sum_{j=1}^{n_b} \mathcal{R}([x_{\ell;j}, (A - \rho_{\ell;j}B)x_{\ell;j}]) = \sum_{j=1}^{n_b} \mathcal{K}_2(A - \rho_{\ell;j}B, x_{\ell;j}).$$

Naturally, as before, to expand search space, $\mathcal{R}([X_\ell, R_\ell])$ through extending each $\mathcal{K}_2(Ax_{\ell;j} - \rho_{\ell;j}B, x_{\ell;j})$ to a high order one. The new extended search subspace now is

$$\sum_{j=1}^{n_b} \mathcal{K}_m(A - \rho_{\ell;j}B, x_{\ell;j}) = \text{span}\{X_\ell, \mathscr{R}_\ell(X_\ell), \ldots, \mathscr{R}_\ell^{m-1}(X_\ell)\} =: \mathcal{K}_m(\mathscr{R}_\ell, X_\ell),$$

where the linear operator $\mathscr{R}_\ell : X \in \mathbb{C}^{n\times n_b} \to \mathscr{R}_\ell(X) = AX - BX\Omega_\ell \in \mathbb{C}^{n\times n_b}$.
$\mathscr{R}_\ell^i(\cdot) = \mathscr{R}_\ell^{i-1}(\mathscr{R}_\ell(\cdot))$, e.g., $\mathscr{R}_\ell^2(X) = \mathscr{R}_\ell(\mathscr{R}_\ell(X))$.

Block Extended SD: make Z basis matrix of $\mathcal{K}_m(\mathscr{R}_\ell, X_\ell)$.

Incorporate Preconditioners

In light of extensive discussions on preconditioning, natural to modify the search subspace to

$$\sum_{j=1}^{n_b} \mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j}),$$

where $K_{\ell;j}$ is the preconditioner intended to move $(\rho_{\ell;j}, x_{\ell;j})$ towards (λ_j, u_j) faster for each j.

Two ways to construct $K_{\ell;j}$:

- $K_{\ell;j} \approx (A - \tilde{\rho}_{\ell;j}B)^{-1}$ for some $\tilde{\rho}_{\ell;j} \neq \rho_{\ell;j}$, ideally closer to λ_j than to any other eigenvalue of $A - \lambda B$.

  Since the eigenvalues of $A - \lambda B$ are unknown, practically make $\tilde{\rho}_{\ell;j}$ closer but not equal to $\rho_{\ell;j}$ than to any other $\rho_{\ell;k}$.

- Perform incomplete LDL^{H} factorization: $A - \rho_{\ell;j}B \approx L_{\ell;j}D_{\ell;j}L_{\ell;j}^{\mathsf{H}}$, where "$\approx$" includes not only the usual "aproximately equal", but also the case when $(A - \rho_{\ell;j}B) - L_{\ell;j}D_{\ell;j}L_{\ell;j}^{\mathsf{H}}$ is approximately a low rank matrix, and $D_{\ell;j} = \mathrm{diag}(\pm 1)$.

  Finally, $K_{i:j} = L_{\ell;j}L_{\ell;j}^{\mathsf{H}}$.

Block Preconditioned Extended SD

Block Preconditioned Extended Steepest Descent method

Given an initial approximation $X_0 \in \mathbb{C}^{n\times n_b}$ with $\mathrm{rank}(X_0) = n_b$, and an integer $m \geq 2$, the algorithm attempts to compute approximate eigenpair to (λ_j, u_j) for $1 \leq j \leq n_b$.

1: compute the eigen-decomposition: $(X_0^{\mathsf{H}} A X_0) W = (X_0^{\mathsf{H}} B X_0) W \Omega_0$, where $W^{\mathsf{H}}(X_0^{\mathsf{H}} B X_0) W = I$, $\Omega_0 = \mathrm{diag}(\rho_{0;1}, \rho_{0;2}, \ldots, \rho_{0;n_b})$;
2: $X_0 = X_0 W$;
3: **for** $\ell = 0, 1, \ldots$ **do**
4: test convergence and lock up the converged (detail to come later);
5: construct preconditioners $K_{\ell;j}$ for $1 \leq j \leq n_b$;
6: compute a basis matrix $Z \in \mathbb{C}^{n\times m n_b}$ of $\sum_{j=1}^{n_b} \mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j} B), x_{\ell;j})$;
7: compute the n_b smallest eigenvalues and corresponding eigenvectors of $Z^{\mathsf{H}}(A - \lambda B)Z$ to get $(Z^{\mathsf{H}} A Z) W = (Z^{\mathsf{H}} B Z) W \Omega_\ell$, where $W^{\mathsf{H}}(Z^{\mathsf{H}} B Z) W = I$, $\Omega_{\ell+1} = \mathrm{diag}(\rho_{\ell+1;1}, \rho_{\ell+1;2}, \ldots, \rho_{\ell+1;n_b})$;
8: $X_{\ell+1} = ZW$;
9: **end for**
10: **return** approximate eigenpairs to (λ_j, u_j) for $1 \leq j \leq n_b$.

Implementation Issue, I

Different preconditioner $K_{\ell;j}$ for each different approximate eigenpair $(\rho_{\ell;j}, x_{\ell;j})$ good for convergence rates, but may not reduce overall time:

- expensive to construct all preconditioners
- cannot compute Z mostly by matrix-matrix multiplications (more later)

Use $K_{\ell;j} \equiv K_\ell$, one preconditioner for all speeding up the convergence of $(\rho_{\ell;1}, x_{\ell;1})$. At the same time other $(\rho_{\ell;j}, x_{\ell;j})$ are making progress, too, but at a slower speed.

Usually $(\rho_{\ell;1}, x_{\ell;1})$ converges first and quickly.

Once $(\rho_{\ell;1}, x_{\ell;1})$ (or the first few in the case of a tight cluster) is determined to be sufficiently accurate, the converged eigenpair is locked up and deflated.

A new preconditioner is computed to aim at the next non-converged eigenpair, and the process continues.

Implementation Issue, II

Need to compute basis matrix $Z \in \mathbb{C}^{n \times mn_b}$ of $\sum_{j=1}^{n_b} \mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j})$.

Z can be gotten by packing the basis matrices of all $\mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j})$ for $1 \leq j \leq n_b$ together. Two drawbacks:

- Such a Z could be ill-conditioned, i.e., columns of Z may not be sufficiently numerically linearly independent; Possible cure: re-orthogonalize packed Z – too costly.
- Building basis for each $\mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j}B), x_{\ell;j})$ uses mostly BLAS-2 operations. Have to be this way if $K_{\ell;j}$ are different.

Different situation if $K_{\ell;j} \equiv K$ (drop iteration step index ℓ). Then

$$\sum_{j=1}^{n_b} \mathcal{K}_m(K(A - \rho_{\ell;j}B), x_{\ell;j}) = \mathcal{K}_m(K\mathscr{R}, X)$$

$$\equiv \text{span}\{X, K\mathscr{R}(X_\ell), \ldots, [K\mathscr{R}]^{m-1}(X)\},$$

where $\mathscr{R}(X) = AX - BX\Omega$, $[K\mathscr{R}]^i(\,\cdot\,) = [K\mathscr{R}]^i(K\mathscr{R}(\,\cdot\,))$, e.g., $[K\mathscr{R}]^2(X) = K\mathscr{R}_\ell(K\mathscr{R}(X))$.

Implementation Issue, II (cont'd)

$Z = [Z_1, Z_2, \ldots, Z_m]$ can be computed by the following block Arnoldi-like process in the B-inner product.

Arnoldi-like process for Z

1: $Z_1 = X$ (recall $X^{\mathsf{H}} B X = I_{n_b}$ already);
2: **for** $i = 2$ to m **do**
3: $\quad Y = K(AZ_{i-1} - B\Omega Z_{i-1})$;
4: $\quad$ **for** $j = 1$ to $i - 1$ **do**
5: $\quad\quad T = Z_j^{\mathsf{H}} B Y$; $Y = Y - Z_j T$;
6: $\quad$ **end for**
7: $\quad Z_i T = Y$ (MGS in the B-inner product);
8: **end for**

Note: At Line 7, Y may not be numerically of full column rank – not a problem.

Anytime if a column is deemed linearly dependent on previous columns, that column should be deleted, along with corresponding ρ_j from Ω.

At completion of MGS, Z_{j+1} will have fewer columns than Y and the size of Ω is shrunk accordingly.

Implementation Issue, III

$(\rho_{\ell;j}, x_{\ell;j})$ is considered acceptable if $\dfrac{\|r_{\ell;j}\|_2}{\|Ax_{\ell;j}\|_2 + |\rho_{\ell;j}| \, \|Bx_{\ell;j}\|_2} \leq \texttt{rtol}$.

Usually λ_j are converged to in order, i.e., the smallest eigenvalues emerge first.

Lock all acceptable approximate eigenpairs in $k_{\text{cvgd}} \times k_{\text{cvgd}}$ diagonal matrix $\boldsymbol{D}$ for eigenvalues and $n \times k_{\text{cvgd}}$ tall matrix $\boldsymbol{U}$ for eigenvectors.

Every time a converged eigenpair is detected, delete the converged $\rho_{\ell;j}$ and $x_{\ell;j}$ from Ω_ℓ and X_ℓ, respectively, and expand $\boldsymbol{D}$ and $\boldsymbol{U}$ to lock up the pair, accordingly.

At the same time, either reduce n_b by 1 or append a (random) B-orthogonal column to X to maintain n_b unchanged.

Deflate to avoid recomputing converged eigenpairs:

1. At Line 7 in the Arnoldi-like process, each column of Z_{j+1} is B-orthogonalized against $\boldsymbol{U}$.
2. Modify $A - \lambda B$ in form, but not explicitly, to $(A + \zeta B\boldsymbol{U}\boldsymbol{U}^{\text{H}}B) - \lambda B$, where ζ should be selected such that $\zeta + \lambda_1 \geq \lambda_{k_{\text{cvgd}}+n_b+1}$. Here we pre-assume the k_{cvgd} converged eigenpairs are indeed those for (λ_j, u_j) for $1 \leq j \leq k_{\text{cvgd}}$. This is usually so, but with no guarantee, of course.

Conjugate Gradient Methods

- Digression: CG for Linear System $Ax = b$
- Conjugate Gradient Method
- Preconditioned Conjugate Gradient Method
- Locally Optimal Conjugate Gradient Method
- Locally Optimal Extended Conjugate Gradient Method
- Locally Optimal Block Preconditioned Extended Conjugate Gradient Method

CG for Linear System $Ax = b$

A is $n \times n$, symmetric, and positive definite. Let

$$\phi(x) = \frac{1}{2} x^{\mathsf{T}} A x - x^{\mathsf{T}} b,$$

quadratic in x, convex, a unique local and global minimum at $x = A^{-1}b$, $\nabla\phi(x) = r(x) \equiv Ax - b$.

CG Algorithm (Hestenes and Stiefel, 1950s):

1. Given x_0, compute $r_0 = Ax_0 - b$, and set $p_0 = -r_0$;
2. For $i = 0, 1, \ldots,$ do

$$\alpha_i = \arg\min_{\alpha} \phi(x_i + \alpha p_i), \quad x_{i+1} = x_i + \alpha_i p_i,$$
$$r_{i+1} = r_i + \alpha_i A p_i, \qquad p_{i+1} = -r_{i+1} + \beta_i p_i.$$

β_i chosen so that $p_{i+1}^{\mathsf{T}} A p_i = 0$; equivalent expressions:

$$\beta_i = \frac{p_i^{\mathsf{T}} A r_{i+1}}{p_i^{\mathsf{T}} A p_i} = \frac{r_{i+1}^{\mathsf{T}} r_{i+1}}{r_i^{\mathsf{T}} r_i} = \frac{r_{i+1}^{\mathsf{T}} (r_{i+1} - r_i)}{r_i^{\mathsf{T}} r_i}.$$

Nonlinear CG

Verbatim translations of Hestenes' and Stiefel's CG to solve

$$\min_x \phi(x), \quad \phi(x) \text{ not necessarily quadratic,}$$

replacing all $r(x_i)$ by $\nabla\phi(x_i)$.

Nonlinear CG Algorithm (Fletcher and Reeves, 1964):

1. Given x_0, compute $\nabla\phi_0 = \nabla\phi(x_0)$, and set $p_0 = -\nabla\phi_0$;
2. For $i = 0, 1, \ldots,$ do

$$\alpha_i = \arg\min_\alpha \phi(x_i + \alpha p_i), \qquad x_{i+1} = x_i + \alpha_i p_i,$$
$$\text{evaluate } \nabla\phi_{i+1} = \nabla\phi(x_{i+1}), \quad p_{i+1} = -\nabla\phi_{i+1} + \beta_i p_i.$$

Several choices for β_i:

$$\beta_i = \frac{\nabla\phi_{i+1}^T \nabla\phi_{i+1}}{\nabla\phi_i^T \nabla\phi_i}, \quad \beta_i = \frac{\nabla\phi_{i+1}^T (\nabla\phi_{i+1} - \nabla\phi_i)}{\nabla\phi_i^T \nabla\phi_i}.$$

Locally Optimal CG

Linear CG: choices of β_i make

- search directions p_i conjugate, i.e., $p_i^T A p_j = 0$ for $i \neq j$.
- CG method terminates in at most n steps.

Nonlinear CG: many nice properties no longer hold for any choice of β_i.

Observe

$$\begin{aligned} x_{i+2} &= x_{i+1} + \alpha_{i+1}(-\nabla\phi_{i+1} + \beta_i p_i) \\ &\in \text{span}\{x_{i+1}, \nabla\phi_{i+1}, p_i\} = \text{span}\{x_{i+1}, \nabla\phi_{i+1}, x_i\}. \end{aligned}$$

Since many nice properties in linear CG are lost anyway in the nonlinear case, why not pick β_i, implicitly, such that (Takahashi, 1965)

$$x_{i+2} = \underset{y \in \text{span}\{x_{i+1}, \nabla\phi_{i+1}, x_i\}}{\arg\min} \phi(y).$$

This gives locally optimal CG. But search over $y \in \text{span}\{x_{i+1}, \nabla\phi_{i+1}, x_i\}$ harder than before.

CG for $Ax = \lambda Bx$

Minimize $\rho(x)$ to compute (λ_1, u_1):

$$\rho(x) = \frac{x^{\mathsf{H}} A x}{x^{\mathsf{H}} B x}, \quad \nabla\rho(x) = \frac{2}{x^{\mathsf{H}} B x} r(x), \quad r(x) := Ax - \rho(x)\, Bx.$$

Line-search $\rho(y) = \inf_{t \in \mathbb{C}} \rho(x + tp)$

1: compute the smaller eigenvalue μ of $X^{\mathsf{H}} A X - \lambda X^{\mathsf{H}} B X$, where $X = [x, p]$, and eigenvector $v = [\nu_1, \nu_2]^T$;

2: $\arg\inf_{t \in \mathbb{C}} \rho(x + tp) =: t_{\text{opt}} = \begin{cases} \nu_2/\nu_1, & \text{if } \nu_1 \neq 0, \\ \infty, & \text{if } \nu_1 = 0; \end{cases}$

3: $y = \begin{cases} x + t_{\text{opt}} p & \text{if } t_{\text{opt}} \text{ is finite,} \\ p & \text{otherwise.} \end{cases}$

CG for $Ax = \lambda Bx$: in nonlinear CG simply replace $\nabla\phi(x)$ by $r(x) := Ax - \rho(x)\, Bx$.

CG for $Ax = \lambda Bx$

CG for $Ax = \lambda Bx$

Given an initial approximation $\boldsymbol{x}_0$ to u_1, and a relative tolerance `rtol`, the algorithm attempts to compute an approximate eigenpair to (λ_1, u_1) with the prescribed `rtol`.

$\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|_B$, $\rho_0 = \boldsymbol{x}_0^{\mathsf{H}} A\boldsymbol{x}_0$, $\boldsymbol{r}_0 = A\boldsymbol{x}_0 - \rho_0 B\boldsymbol{x}_0$, $\boldsymbol{p}_0 = -\boldsymbol{r}_0$;
for $\ell = 0, 1, \ldots$ **do**
  if $\|\boldsymbol{r}_\ell\|_2/(\|A\boldsymbol{x}_\ell\|_2 + |\rho_\ell|\,\|B\boldsymbol{x}_\ell\|_2) \le$ `rtol` **then**
    BREAK;
  else
    compute $\alpha_\ell = t_{\text{opt}} := \inf_{t\in\mathbb{C}} \rho(\boldsymbol{x}_\ell + t\boldsymbol{p}_\ell)$, and then

$$y = \begin{cases} \boldsymbol{x}_\ell + \alpha_\ell \boldsymbol{p}_\ell & \text{if } \alpha_\ell \text{ is finite,} \\ \boldsymbol{p}_\ell & \text{otherwise.} \end{cases}$$

    $\boldsymbol{x}_{\ell+1} = y/\|y\|_B$;
    set $\rho_{\ell+1} = \boldsymbol{x}_{\ell+1}^{\mathsf{H}} A\boldsymbol{x}_{\ell+1}$, $\boldsymbol{r}_{\ell+1} = A\boldsymbol{x}_{\ell+1} - \rho_{\ell+1} B\boldsymbol{x}_{\ell+1}$, $\boldsymbol{p}_{\ell+1} = -\boldsymbol{r}_{\ell+1} + \beta_\ell \boldsymbol{p}_\ell$,
    where $\beta_\ell = \frac{\boldsymbol{r}_{\ell+1}^{\mathsf{H}} \boldsymbol{r}_{\ell+1}}{\boldsymbol{r}_\ell^{\mathsf{H}} \boldsymbol{r}_\ell}$ or $\frac{\boldsymbol{r}_{\ell+1}^{\mathsf{H}} (\boldsymbol{r}_{\ell+1} - \boldsymbol{r}_\ell)}{\boldsymbol{r}_\ell^{\mathsf{H}} \boldsymbol{r}_\ell}$
  end if
end for
return $(\rho_\ell, \boldsymbol{x}_\ell)$ as an approximate eigenpair to (λ_1, u_1).

A Convergence Theorem

Convergence Theorem (Yang, 1993)

With $\beta_\ell = \frac{\boldsymbol{r}_{\ell+1}^{\mathsf{H}} \boldsymbol{r}_{\ell+1}}{\boldsymbol{r}_\ell^{\mathsf{H}} \boldsymbol{r}_\ell}$, $\boldsymbol{\rho}_\ell$ converges to some eigenvalue $\hat{\lambda}$ of $A - \lambda B$ and there is a convergent subsequence $\{\boldsymbol{x}_{\ell_i}\}$ of $\{\boldsymbol{x}_\ell\}$ such that

$$\|(A - \hat{\lambda} B)\boldsymbol{x}_{\ell_i}\|_2 \to 0 \quad \text{as} \quad i \to \infty,$$

i.e., $\boldsymbol{x}_{\ell_i}$ converges in direction to a corresponding eigenvector.

If $\hat{\lambda} = \lambda_1$, then $\|(A - \hat{\lambda} B)\boldsymbol{x}_\ell\|_2 \to 0$ as $\ell \to \infty$, i.e., $\boldsymbol{x}_\ell$ converges in direction to a corresponding eigenvector. (*seem new*)

- First part due to (Yang, 1993); second part seems new.
- Only for $\beta_\ell = \frac{\boldsymbol{r}_{\ell+1}^{\mathsf{H}} \boldsymbol{r}_{\ell+1}}{\boldsymbol{r}_\ell^{\mathsf{H}} \boldsymbol{r}_\ell}$, however.
- Proof much more complicated than for SD.
- Rate of convergence: mostly heuristic, none rigorous proven.

Preconditioned CG for $Ax = \lambda Bx$

As in SD, Preconditioned CG = *vanilla* CG on $L^{-\mathsf{H}}AL^{-1} - \lambda L^{-\mathsf{H}}BL^{-1}$.

Let $\widetilde{A} - \lambda\widetilde{B} := L^{-\mathsf{H}}AL^{-1} - \lambda L^{-\mathsf{H}}BL^{-1}$. Adopt notation convention for $\widetilde{A} - \lambda\widetilde{B}$: same symbols but with *tilde*s. E.g., $\widetilde{x} = Lx$,

$$\widetilde{\rho}(\widetilde{x}) = \frac{\widetilde{x}^{\mathsf{H}}L^{-\mathsf{H}}AL^{-1}\widetilde{x}}{\widetilde{x}^{\mathsf{H}}L^{-\mathsf{H}}BL^{-1}\widetilde{x}} \equiv \rho(x), \quad \widetilde{r}(\widetilde{x}) = L^{-\mathsf{H}}AL^{-1}\widetilde{x} - \widetilde{\rho}(\widetilde{x})L^{-\mathsf{H}}BL^{-1}\widetilde{x} \equiv L^{-\mathsf{H}}r(x).$$

Key CG step:

$$\widetilde{\alpha}_\ell = \arg\min_{\widetilde{\alpha}} \widetilde{\rho}(\widetilde{\boldsymbol{x}}_\ell + \widetilde{\alpha}\widetilde{\boldsymbol{p}}_\ell), \qquad \widetilde{\boldsymbol{x}}_{\ell+1} = \widetilde{\boldsymbol{x}}_\ell + \widetilde{\alpha}_\ell\widetilde{\boldsymbol{p}}_\ell,$$

$$\widetilde{\boldsymbol{r}}_{\ell+1} = L^{-\mathsf{H}}AL^{-1}\widetilde{\boldsymbol{x}}_{\ell+1} - \widetilde{\rho}(\widetilde{\boldsymbol{x}}_{\ell+1})L^{-\mathsf{H}}BL^{-1}\widetilde{\boldsymbol{x}}_{\ell+1}, \quad \widetilde{\boldsymbol{p}}_{\ell+1} = -\widetilde{\boldsymbol{r}}_{\ell+1} + \widetilde{\beta}_\ell\widetilde{\boldsymbol{p}}_\ell.$$

Perform substitutions $\widetilde{\boldsymbol{x}}_\ell = L\boldsymbol{x}_\ell$ and $\widetilde{\boldsymbol{r}}_\ell = L^{-\mathsf{H}}\boldsymbol{r}_\ell$:

$$\widetilde{\alpha}_\ell = \arg\min_{\widetilde{\alpha}} \rho(\boldsymbol{x}_\ell + \widetilde{\alpha}\underbrace{L^{-1}\widetilde{\boldsymbol{p}}_\ell}_{=:\boldsymbol{p}_\ell}), \qquad \boldsymbol{x}_{\ell+1} = \boldsymbol{x}_\ell + \widetilde{\alpha}_\ell L^{-1}\widetilde{\boldsymbol{p}}_\ell,$$

$$\boldsymbol{r}_{\ell+1} = A\boldsymbol{x}_{\ell+1} - \rho(\boldsymbol{x}_{\ell+1})B\boldsymbol{x}_{\ell+1}, \quad \underbrace{L^{-1}\widetilde{\boldsymbol{p}}_{\ell+1}}_{=:\boldsymbol{p}_{\ell+1}} = -\underbrace{(L^{\mathsf{H}}L)^{-1}}_{=:K}\boldsymbol{r}_{\ell+1} + \widetilde{\beta}_\ell\underbrace{L^{-1}\widetilde{\boldsymbol{p}}_\ell}_{=:\boldsymbol{p}_\ell}.$$

Preconditioned CG for $Ax = \lambda Bx$

Preconditioned CG for $Ax = \lambda Bx$

Given an initial approximation $\boldsymbol{x}_0$ to u_1, a (positive definite) preconditioner K, and a relative tolerance `rtol`, the algorithm attempts to compute an approximate pair to (λ_1, u_1) with the prescribed `rtol`.

1: $\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|_B$, $\rho_0 = \boldsymbol{x}_0^{\mathsf{H}} A\boldsymbol{x}_0$, $\boldsymbol{r}_0 = A\boldsymbol{x}_0 - \rho_0 B\boldsymbol{x}_0$, $\boldsymbol{p}_0 = -K\boldsymbol{r}_0$;
2: **for** $\ell = 0, 1, \ldots$ **do**
3: **if** $\|\boldsymbol{r}_\ell\|_2/(\|A\boldsymbol{x}_\ell\|_2 + |\rho_\ell|\,\|B\boldsymbol{x}_\ell\|_2) \le$ `rtol` **then**
4: **BREAK**;
5: **else**
6: compute $\alpha_\ell = t_{\text{opt}} := \inf_{t\in\mathbb{C}} \rho(\boldsymbol{x}_\ell + t\boldsymbol{p}_\ell)$, and then

$$y = \begin{cases} \boldsymbol{x}_\ell + \alpha_\ell \boldsymbol{p}_\ell & \text{if } \alpha_\ell \text{ is finite,} \\ \boldsymbol{p}_\ell & \text{otherwise.} \end{cases}$$

7: $\boldsymbol{x}_{\ell+1} = y/\|y\|_B$;
8: set $\rho_{\ell+1} = \boldsymbol{x}_{\ell+1}^{\mathsf{H}} A\boldsymbol{x}_{\ell+1}$, $\boldsymbol{r}_{\ell+1} = A\boldsymbol{x}_{\ell+1} - \rho_{\ell+1} B\boldsymbol{x}_{\ell+1}$,

$$\boldsymbol{p}_{\ell+1} = -K\boldsymbol{r}_{\ell+1} + \beta_\ell \boldsymbol{p}_\ell, \text{ where } \beta_\ell = \frac{\boldsymbol{r}_{\ell+1}^{\mathsf{H}} K \boldsymbol{r}_{\ell+1}}{\boldsymbol{r}_\ell^{\mathsf{H}} K \boldsymbol{r}_\ell} \text{ or } \frac{\boldsymbol{r}_{\ell+1}^{\mathsf{H}} K (\boldsymbol{r}_{\ell+1} - \boldsymbol{r}_\ell)}{\boldsymbol{r}_\ell^{\mathsf{H}} K \boldsymbol{r}_\ell}.$$

9: **end if**
10: **end for**
11: **return** $(\rho_\ell, \boldsymbol{x}_\ell)$ as an approximate eigenpair to (λ_1, u_1).

Convergence

Earlier discussions on selecting a good preconditioner for PSD should apply:

- $A - \sigma B = LDL^{\mathsf{H}}$, $D = \mathrm{diag}(\pm 1)$, $K = (L^{\mathsf{H}}L)^{-1}$.

Various heuristics on the convergence rates of the preconditioned CG, but none is rigorously proved. Even less can be said about the theoretical analysis of block (or subspace) versions of the preconditioned CG method (to come soon).

But since preconditioned CG is CG for $L^{-\mathsf{H}}AL^{-1} - \lambda L^{-\mathsf{H}}BL^{-1}$, previous convergence theorem for CG remains valid.

Locally Optimal CG for $Ax = \lambda Bx$

In writing down CG for $Ax = \lambda Bx$, we did

- gradient-to-residual replacement: replacing the gradient by the eigen-residual $r(x) = Ax - \rho(x)Bx$ which differs by a scalar factor $2/x^{\mathsf{H}}Bx$ from the gradient $\nabla\rho(x) = \frac{2}{x^{\mathsf{H}}Bx}\,[Ax - \rho(x)\,Bx]$;
- also normalizing $\boldsymbol{x}_\ell$. No theory around as to why we should normalize $\boldsymbol{x}_\ell$, beside that they are some eigenvector approximations.

We made a couple of "arbitrary choices". Their effects on the rate of convergence are not clear.

Locally optimal CG eliminates the "arbitrariness" altogether: compute $\boldsymbol{x}_{\ell+1}$ from the subspace $\operatorname{span}\{\boldsymbol{x}_{\ell-1}, \boldsymbol{x}_\ell, \boldsymbol{r}_\ell\}$ by

$$\min_{x \in \operatorname{span}\{\boldsymbol{x}_{\ell-1}, \boldsymbol{x}_\ell, \boldsymbol{r}_\ell\}} \rho(x),$$

which is solvable through the Rayleigh-Ritz procedure.

Locally Optimal CG

Locally Optimal CG for $Ax = \lambda Bx$

Given an initial approximation $\boldsymbol{x}_0$ to u_1, and a relative tolerance `rtol`, the algorithm attempts to compute an approximate eigenpair to (λ_1, u_1) with the prescribed `rtol`.

1: $\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|_B$, $\boldsymbol{\rho}_0 = \boldsymbol{x}_0^{\mathsf{H}} A\boldsymbol{x}_0$, $\boldsymbol{r}_0 = A\boldsymbol{x}_0 - \boldsymbol{\rho}_0 B\boldsymbol{x}_0$, $\boldsymbol{x}_{-1} = 0$;
2: **for** $\ell = 0, 1, \ldots$ **do**
3: **if** $\|\boldsymbol{r}_\ell\|_2/(\|A\boldsymbol{x}_\ell\|_2 + |\boldsymbol{\rho}_\ell|\,\|B\boldsymbol{x}_\ell\|_2) \leq$ `rtol` **then**
4: **BREAK**;
5: **else**
6: compute a basis matrix $Z \in \mathbb{C}^{n\times k}$ ($k = 2$ or 3) of the subspace $\mathrm{span}\{\boldsymbol{x}_\ell, \boldsymbol{x}_{\ell-1}, \boldsymbol{r}_\ell\}$;
7: compute the smallest eigenvalue μ and corresponding eigenvector v of $Z^{\mathsf{H}}(A - \lambda B)Z$;
8: $y = Zv$, $\boldsymbol{x}_{\ell+1} = y/\|y\|_B$;
9: $\boldsymbol{\rho}_{\ell+1} = \mu$, $\boldsymbol{r}_{\ell+1} = A\boldsymbol{x}_{\ell+1} - \boldsymbol{\rho}_{\ell+1} B\boldsymbol{x}_{\ell+1}$;
10: **end if**
11: **end for**
12: **return** $(\boldsymbol{\rho}_\ell, \boldsymbol{x}_\ell)$ as an approximate eigenpair to (λ_1, u_1).

Implementation

$\boldsymbol{x}_\ell$ moves closer and closer to u_1; $\boldsymbol{x}_\ell$, $\boldsymbol{x}_{\ell-1}$ increasingly move towards being linearly dependent.

Line 6: Z contaminated more and more by rounding errors. How to mitigate that?

To replace $\boldsymbol{x}_{\ell-1}$ by some $\boldsymbol{y}_\ell := \xi_{\ell,1}\boldsymbol{x}_\ell - \xi_{\ell,2}\boldsymbol{x}_{\ell-1}$ such that

$$\text{span}\{\boldsymbol{x}_\ell, \boldsymbol{x}_{\ell-1}, \boldsymbol{r}_\ell\} = \text{span}\{\boldsymbol{x}_\ell, \boldsymbol{y}_\ell, \boldsymbol{r}_\ell\}.$$

Then same (μ, ν) at Line 7. But need to generate $\boldsymbol{y}_{\ell+1}$, given $\boldsymbol{x}_\ell$, $\boldsymbol{y}_\ell$, $\boldsymbol{r}_\ell$.

$Z = [z_1, z_2, z_3]$ is B-orthonormal (by MGS), and $z_1 = \boldsymbol{x}_\ell$. Then
$y = Zv = \nu_1 z_1 + \nu_2 z_2 + \nu_3 z_3 = \nu_1 \boldsymbol{x}_\ell + \nu_2 z_2 + \nu_3 z_3$.

Set $\boldsymbol{y}_{\ell+1} := y - \nu_1 \boldsymbol{x}_\ell = \|y\|_B \boldsymbol{x}_{\ell+1} - \nu_1 \boldsymbol{x}_\ell =: \xi_{\ell+1,1}\boldsymbol{x}_{\ell+1} - \xi_{\ell+1,2}\boldsymbol{x}_\ell$.

Modify Lines 1, 6, and 8 as follows while keeping others the same.

1: $\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|_B$, $\boldsymbol{\rho}_0 = \boldsymbol{x}_0^{\mathsf{H}} A\boldsymbol{x}_0$, $\boldsymbol{r}_0 = A\boldsymbol{x}_0 - \boldsymbol{\rho}_0 B\boldsymbol{x}_0$, $\boldsymbol{y}_0 = 0$;
6: compute a basis matrix $Z \in \mathbb{C}^{n\times k}$ ($k = 2$ or 3) of the subspace $\text{span}\{\boldsymbol{x}_\ell, \boldsymbol{y}_\ell, \boldsymbol{r}_\ell\}$;
8: $y = Zv$, $\boldsymbol{x}_{\ell+1} = y/\|y\|_B$, $\boldsymbol{y}_{\ell+1} = Z\hat{v}$, where $\hat{v}$ is v with its 1st entry zeroed;

Convergence

Convergence Theorem for LOCG)

$\boldsymbol{\rho}_\ell$ converges to some eigenvalue $\hat{\lambda}$ of $A - \lambda B$ and $\|(A - \hat{\lambda}B)\boldsymbol{x}_\ell\|_2 \to 0$ as $\ell \to \infty$, i.e., $\boldsymbol{x}_\ell$ converges in direction to a corresponding eigenvector.

- Same convergence theorem for SD;
- For CG, PCG, only $\|(A - \hat{\lambda}B)\boldsymbol{x}_{\ell_i}\|_2 \to 0$;
- Inclusion of the residual $\boldsymbol{r}_\ell$ makes the difference.

Improve LOCG

Three ideas for improving SD naturally apply here:

1. Incorporate a preconditioner K: simply modify $\boldsymbol{r}_\ell$ to $K\boldsymbol{r}_\ell$;
2. Extend search space from currently

$$\text{span}\{\boldsymbol{x}_{\ell-1}\} + \mathcal{K}_2(A - \rho_\ell B, \boldsymbol{x}_\ell) \quad \text{to} \quad \text{span}\{\boldsymbol{x}_{\ell-1}\} + \mathcal{K}_m(A - \rho_\ell B, \boldsymbol{x}_\ell);$$

3. Use block $X_0 \in \mathbb{C}^{n \times n_b}$.

The ideas can be applied in any combination ($2^3 = 8$ of them): E.g.,

- Locally Optimal Preconditioned CG (LOPCG): $m = 2$, $n_b = 1$, $K \neq I$;
- Locally Optimal Block Preconditioned CG (LOBPCG): $m = 2$, $n_b > 1$, $K \neq I$;
- Locally Optimal Extended CG (LOECG): $m > 2$, $n_b = 1$, $K = I$;
- Locally Optimal Preconditioned Extended CG (LOPECG): $m > 2$, $n_b = 1$, $K \neq I$;
- Locally Block Optimal Preconditioned Extended CG (LOBPECG): $m > 2$, $n_b > 1$, $K \neq I$.

Locally Optimal CG

Extended Locally Block Optimal Preconditioned CG

Given an initial approximation $X_0 \in \mathbb{C}^{n\times n_b}$ with $\mathrm{rank}(X_0) = n_b$, and an integer $m \geq 2$, the algorithm attempts to compute approximate eigenpairs to (λ_j, u_j) for $1 \leq j \leq n_b$.

1: compute the eigen-decomposition: $(X_0^{\mathsf{H}} A X_0) W = (X_0^{\mathsf{H}} B X_0) W \Omega_0$, where $W^{\mathsf{H}}(X_0^{\mathsf{H}} B X_0) W = I$, $\Omega_0 = \mathrm{diag}(\rho_{0;1}, \rho_{0;2}, \ldots, \rho_{0;n_b})$;
2: $X_0 = X_0 W$, and $X_{-1} = 0$;
3: **for** $\ell = 0, 1, \ldots$ **do**
4: test convergence and lock up the converged (detail as in EBPSD);
5: construct preconditioners $K_{\ell;j}$ for $1 \leq j \leq n_b$;
6: compute a basis matrix $Z \in \mathbb{C}^{n\times(m+1)n_b}$ of the subspace
$$\sum_{j=1}^{n_b} \mathcal{K}_m(K_{\ell;j}(A - \rho_{\ell;j} B), x_{\ell;j}) + \mathcal{R}(X_{\ell-1});$$
7: compute the n_b smallest eigenvalues and corresponding eigenvectors of $Z^{\mathsf{H}}(A - \lambda B)Z$ to get $(Z^{\mathsf{H}} A Z) W = (Z^{\mathsf{H}} B Z) W \Omega_\ell$, where $W^{\mathsf{H}}(Z^{\mathsf{H}} B Z) W = I$, $\Omega_{\ell+1} = \mathrm{diag}(\rho_{\ell+1;1}, \rho_{\ell+1;2}, \ldots, \rho_{\ell+1;n_b})$;
8: $X_{\ell+1} = ZW$;
9: **end for**
10: **return** approximate eigenpairs to (λ_j, u_j) for $1 \leq j \leq n_b$.

Implementation

Three important implementation issues earlier for XBPSD essentially apply here, but more need to be said about Z at Line 6 here.

$X_{\ell-1}$ can be replaced by something else, using the idea earlier for LOCG. Specifically, Lines 2, 6, and 8 should be modified to

---

2: $X_0 = X_0 W$, and $Y_0 = 0$;

6: compute a basis matrix $Z \in \mathbb{C}^{n\times(m+1)n_b}$ of the subspace $\sum_{j=1}^{n_b} \mathcal{K}_m(K_{\ell;j}(A-\rho_{\ell;j}B), x_{\ell;j}) + \mathcal{R}(Y_\ell)$ such that $\mathcal{R}(Z_{(:,1:n_b)}) = \mathcal{R}(X_\ell)$;

8: $X_{\ell+1} = ZW$, $Y_{\ell+1} = Z\widehat{W}$, where $\widehat{W}$ is W with its n_b rows zeroed;

For $K_{\ell;j} \equiv K_\ell$, Z is basis matrix of (dropping the subscript ℓ)

$$\mathcal{K}_m(K\mathscr{R}, X) + \mathcal{R}(Y) = \text{span}\{X, K\mathscr{R}(X), \ldots, [K\mathscr{R}]^{m-1}(X)\} + \mathcal{R}(Y).$$

1. compute a basis matrix $[Z_1, Z_2, \ldots, Z_m]$ for $\mathcal{K}_m(K\mathscr{R}, X)$ by the Block Arnoldi-like process in the B-inner product. In particular, $Z_1 = X$.
2. B-orthogonalize Y against $[Z_1, Z_2, \ldots, Z_m]$ to get Z_{m+1} satisfying $Z_{m+1}^{\mathsf{H}} B Z_{m+1} = I$.
3. $Z = [Z_1, Z_2, \ldots, Z_{m+1}]$.

References

Precise rates of convergence for various CG methods are scarce and not well understood, especially so for methods of block version. The existing research on the convergence of various SD and CG-type methods, although fragmental and incomplete, should be helpful and provide heuristic insights. Some of the references are


L. Bergamaschi, G. Gambolati, and G. Pini. Asymptotic convergence of conjugate gradient methods for the partial symmetric eigenproblem. *Numer. Linear Algebra Appl.*, 4(2):69–84, 1997.


J. H. Bramble, J. E. Pasciak, and A. V. Knyazev. A subspace preconditioning algorithm for eigenvector/eigenvalue computation. *Adv. in Comput. Math.*, 6:159–189, 1996.


Andrew V. Knyazev and Klaus Neymeyr. A geometric theory for preconditioned inverse iteration III: A short and sharp convergence estimate for generalized eigenvalue problems. *Linear Algebra Appl.*, 358(1-3):95–114, 2003.


D. E. Longsine and S. F. McCormick. Simultaneous Rayleigh-quotient minimization methods for $Ax = \lambda Bx$. *Linear Algebra Appl.*, 34:195–234, 1980.


S. Oliveira. On the convergence rate of a preconditioned subspace eigensolver. *Computing*, 63:219–231, 1999.


E. E. Ovtchinnikov. Jacobi correction equation, line search, and conjugate gradients in hermitian eigenvalue computation I: Computing an extreme eigenvalue. *SIAM J. Numer. Anal.*, 46(5):2567–2592, 2008.


E. E. Ovtchinnikov. Jacobi correction equation, line search, and conjugate gradients in hermitian eigenvalue computation II: Computing several extreme eigenvalues. *SIAM J. Numer. Anal.*, 46(5):2593–2619, 2008.


H. Yang. Conjugate gradient methods for the Rayleigh quotient minimization of generalized eigenvalue problems. *Computing*, 51:79–94, 1993.

Extending Min-Max Principles: Indefinite B

- Early Extensions
- Positive Semi-definite Pencil

Early Extensions

Min-max principles, Cauchy interlace inequalities are Foundations for optimization approaches to solve few extreme eigenpairs of $A - \lambda B$ with $B \succ 0$.

How far can these theoretical results be extended?

Early extensions (before 1982) of Courant-Fischer min-max princples:

- $Ax = \lambda x$ with $A \succeq 0$ in an indefinite inner product

  R. S. Phillips. A minimax characterization for the eigenvalues of a positive symmetric operator in a space with an indefinite metric. *J. Fac. Sci. Univ. Tokyo Sect. I*, 17:51–59, 1970.

  B. Textorius. Minimaxprinzipe zur Bestimmung der Eigenwerte J-nichtnegativer Operatoren. *Math. Scand.*, 35:105–114, 1974.

  It turns out to be a special case of $A - \lambda B$ with indefinite and nonsingular B.

- Hyperbolic $Q(\lambda) = A\lambda^2 + B\lambda + C$:

  R. Duffin. A minimax theory for overdamped networks. *Indiana Univ. Math. J.*, 4:221–233, 1955.

- More general nonlinear eigenvalue problems:

  H. Voss and B. Werner. A minimax principle for nonlinear eigenvalue problems with applications to nonoverdamped systems. *Math. Meth. Appl. Sci.*, 4:415–424, 1982.

  and references therein.

Will focus on $A - \lambda B$ with indefinite B and hyperbolic $Q(\lambda)$ ...

Hermitian $Ax = \lambda Bx$ with Indefinite B

Difficulties:

- B is indefinite — $Ax = \lambda Bx$ not equivalent to standard Hermitian eigenvalue problem;
- $Ax = \lambda Bx$ may have complex eigenvalues — no min-max for complex eigenvalues;

  The case we consider will turn out to have only real eigenvalues.
- $A - \lambda B$ may be a singular pencil: $\det(A - \lambda B) \equiv 0$ for $\lambda \in \mathbb{C}$. Need a definition for eigenvalues.

History

1. Extending Courant-Fischer for regular Hermitian pencil:
   - Nonsingular B: Lancaster & Ye (1989), Ye's thesis (1989), Najman & Ye (1991), Binding & Ye (1995)
   - Singular B: Najman & Ye (1993), Binding & Najman & Ye (1999)

   Only certain real semi-simple eigenvalues admit a Courant-Fischer type characterization.

2. Extending trace min for positive semi-definite Hermitian pencil (i.e., $A - \lambda_0 B \succeq 0$ for some $\lambda_0 \in \mathbb{R}$):
   - Nonsingular B: Kovač-Striko & Veselić (1995)
   - (Possibly) singular pencil $A - \lambda B$: Liang & Li & Bai (2012)

3. Extending Wielandt's min-max for positive semi-definite Hermitian pencil:
   - Nonsingular B: Nakić and Veselić (2003) (actually for regular Hermitian $A - \lambda B$, but beware of inaccurate/incorrect statements/equations there)
   - (Possibly) singular pencil $A - \lambda B$: Liang & Li (2012)

4. Extending trace min for linear response eigenvalue problem: Bai & Li (2011).

Positive Semi-definite Pencil

$A = A^{\mathsf{H}}$, $B = B^{\mathsf{H}} \in \mathbb{C}^{n\times n}$.

1. Positive semi-definite pencil: $A - \lambda_0 B \succeq 0$ for some $\lambda_0 \in \mathbb{R}$;
   $A - \lambda B$ will be assumed so hereafter.
2. Finite eigenvalue $\mu(\neq \infty)$: $\mathrm{rank}(A - \mu B) < \max_{\lambda\in\mathbb{C}} \mathrm{rank}(A - \lambda B)$; This allows singular pencil $A - \lambda B$.
3. Eigenvector x: $Ax = \mu Bx$ and $x \notin \mathcal{N}(A) \cap \mathcal{N}(B)$ and .
4. B's Inertia (n_+, n_0, n_-): n_+ positive, n_0 zero, and n_- negative eigenvalues, respectively.
5. Can prove: *positive semi-definite pencil $A - \lambda B$ has only*

   $$r := \mathrm{rank}(B) = n_+ + n_-$$

   finite eigenvalues all of which are real:

   $$\lambda_{n_-}^- \le \cdots \le \lambda_1^- \le \lambda_0 \le \lambda_1^+ \le \cdots \le \lambda_{n_+}^+.$$

6. In more detail ...

Canonical Form of Positive Semi-definite Pencil

1. There exists a nonsingular $W \in \mathbb{C}^{n\times n}$ such that

$$W^{\mathsf{H}}AW = \begin{array}{c} \\ n_1 \\ r-n_1 \\ n-r \end{array}\!\!\begin{array}{c} \begin{array}{ccc} n_1 & r-n_1 & n-r \end{array} \\ \begin{bmatrix} \Lambda_1 & & \\ & \Lambda_0 & \\ & & \Lambda_\infty \end{bmatrix} \end{array}, \quad W^{\mathsf{H}}BW = \begin{array}{c} \\ n_1 \\ r-n_1 \\ n-r \end{array}\!\!\begin{array}{c} \begin{array}{ccc} n_1 & r-n_1 & n-r \end{array} \\ \begin{bmatrix} \Omega_1 & & \\ & \Omega_0 & \\ & & 0 \end{bmatrix} \end{array},$$

   - $\Lambda_1 = \mathrm{diag}(s_1\alpha_1, \ldots, s_\ell\alpha_\ell)$, $\Omega_1 = \mathrm{diag}(s_1, \ldots, s_\ell)$, $s_i = \pm 1$, and $\Lambda_1 - \lambda_0\Omega_1 \succ 0$;
   - $\Lambda_0 = \mathrm{diag}(\Lambda_{0,1}, \ldots, \Lambda_{0,m+m_0})$, $\Omega_0 = \mathrm{diag}(\Omega_{0,1}, \ldots, \Omega_{0,m+m_0})$,

$$\Lambda_{0,i} = t_i\lambda_0, \qquad \Omega_{0,i} = t_i = \pm 1, \quad \text{for } 1 \le i \le m,$$
$$\Lambda_{0,i} = \begin{bmatrix} 0 & \lambda_0 \\ \lambda_0 & 1 \end{bmatrix}, \ \Omega_{0,i} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{for } m+1 \le i \le m+m_0.$$

     There are no such pair (Λ_0, Ω_0) if $A - \lambda_0 B \succ 0$.
   - $\Lambda_\infty = \mathrm{diag}(\alpha_{r+1}, \ldots, \alpha_n) \succeq 0$ with $\alpha_i \in \{1, 0\}$ for $r+1 \le i \le n$.

2. $A - \lambda B$ has $n_+ + n_-$ finite eigenvalues all of which are real. Denote these finite eigenvalues by $\lambda_i^{\pm}$ and arrange them

$$\lambda_{n_-}^- \le \cdots \le \lambda_1^- \le \lambda_1^+ \le \cdots \le \lambda_{n_+}^+.$$

3. $\{\gamma \in \mathbb{R} \mid A - \gamma B \succeq 0\} = [\lambda_{n_-}^-, \lambda_1^+]$. Moreover, if $A - \lambda B$ is regular, then $A - \lambda B$ is a positive definite pencil if and only if $\lambda_{n_-}^- < \lambda_1^+$, in which case $\{\gamma \in \mathbb{R} \mid A - \gamma B \succ 0\} = (\lambda_{n_-}^-, \lambda_1^+)$.

(Proof in Liang, Li, & Bai, *Linear Algebra and its Applications*, **438** (2013), 3085-3106)

Courant-Fischer Type

Courant-Fischer type min-max principle

$$\lambda_i^+ = \sup_{\operatorname{codim} \mathcal{X} = i-1} \inf_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x > 0}} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x}, \quad \lambda_i^+ = \inf_{\dim \mathcal{X} = i} \sup_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x > 0}} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x} \quad \text{for } 1 \le i \le n_+,$$

$$\lambda_i^- = \inf_{\operatorname{codim} \mathcal{X} = i-1} \sup_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x < 0}} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x}, \quad \lambda_i^- = \sup_{\dim \mathcal{X} = i} \inf_{\substack{x \in \mathcal{X} \\ x^{\mathrm{H}} B x < 0}} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x} \quad \text{for } 1 \le i \le n_-.$$

In particular,
$$\lambda_1^+ = \inf_{x^{\mathrm{H}} B x > 0} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x}, \quad \lambda_1^- = \sup_{x^{\mathrm{H}} B x < 0} \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} B x}.$$

- Lancaster & Ye (1989), Ye's thesis (1989) for diagonalizable $A - \lambda B$ and B nonsingular. (Actually studied $A - \lambda B$ not necessarily positive semi-definite, but then only some of the eigenvalues can be characterized.)
- Najman & Ye (1993), Binding & Najman & Ye (1999) for regular $A - \lambda B$. (Actually studied $A - \lambda B$ not necessarily positive semi-definite, but then only some of the real eigenvalues can be characterized.)
- Liang & Li (2012) for allowing singular pencil $A - \lambda B$.

Trace Min Type

$$k_+ \le n_+,\ k_- \le n_-,\ k := k_+ + k_- \ge 1, \quad J_k = \begin{bmatrix} I_{k_+} & \\ & -I_{k_-} \end{bmatrix}.$$

Trace minimization principle

$$\inf_{\substack{X_+=[x_1,\ldots,x_{k_+}] \\ X_-=[y_1,\ldots,y_{k_-}] \\ X=[X_+,X_-], \\ \text{subject to (8)}}} \operatorname{trace}(X^{\mathsf{H}}AX) = \sum_{i=1}^{k_+} \lambda_i^+ - \sum_{i=1}^{k_-} \lambda_i^-.$$

$$\text{either } X^{\mathsf{H}}BX = J_k, \text{ or } X_+^{\mathsf{H}}BX_+ = I_{k_+} \text{ and } X_-^{\mathsf{H}}BX_- = -I_{k_-}. \tag{8}$$

A converse: $\inf_{X^{\mathsf{H}}BX=J_k} \operatorname{trace}(X^{\mathsf{H}}AX) > -\infty \Rightarrow A - \lambda B$ positive semi-definite.

- Kovač-Striko & Veselić (1995) for B nonsingular, subject to $X^{\mathsf{H}}BX = J_k$.
- Liang & Li & Bai (2012) for allowing singular pencil $A - \lambda B$.

Unfortunately no Trace Max in general.

Cauchy Type

Eigenvalues of $A - \lambda B$: $\lambda_{n_-}^- \le \cdots \le \lambda_1^- \le \lambda_1^+ \le \cdots \le \lambda_{n_+}^+$.

$$k_+ \le n_+,\ k_- \le n_-,\ k := k_+ + k_- \ge 1, \quad J_k = \begin{bmatrix} I_{k_+} & \\ & -I_{k_-} \end{bmatrix};$$

$X \in \mathbb{C}^{k\times k}$, $X^{\mathrm{H}} B X = J_k$, or the inertia of $X^{\mathrm{H}} B X$ is $(k_+, 0, k_-)$;

Eigenvalues of $X^{\mathrm{H}}(A - \lambda B)X$: $\mu_{k_-}^- \le \cdots \le \mu_1^- \le \mu_1^+ \le \cdots \le \mu_{k_+}^+$.

Cauchy-type interlacing inequality

$$\begin{aligned} \lambda_i^+ \le \mu_i^+ \le \lambda_{i+n-k}^+, &\quad \text{for } 1 \le i \le k_+, \\ \lambda_{j+n-k}^- \le \mu_i^- \le \lambda_i^-, &\quad \text{for } 1 \le j \le k_-, \end{aligned}$$

where undefined $\lambda_i^+ = \infty$ for $i > n_+$ and undefined $\lambda_j^- = -\infty$ for $j > n_-$.

- Kovač-Striko & Veselić (1995) for B nonsingular.
- Liang & Li & Bai (2012) for allowing singular pencil $A - \lambda B$.

These results potentially lead to optimization approaches to compute 1st few $\lambda_i^{\pm}$ (these are interior eigenvalues!). See Bai & Li (2011, 2012, 2013 for linear response eigenvalue problem), Kressner, Pandur, & Shao (2013).

Linear Response Eigenvalue Problem

- Background
- Basic Theory
- Minimization Principles
- 4D SD and 4D CG type Methods

DFT, strictly a ground-state theory, cannot be applied to study the excitations of systems that are involved in Optical Absorption Spectra (OAS) calculations.

Runge and Gross (1984) generalized DFT to Time-Dependent Density Functional Theory (TD-DFT):

$$\iota\frac{\partial}{\partial t}\phi_i(\boldsymbol{r},t) = \Big[-\frac{1}{2}\nabla^2 + \underbrace{\int \frac{n(\boldsymbol{r}',t)}{|\boldsymbol{r}-\boldsymbol{r}'|}\,d\boldsymbol{r}' + \frac{\delta A_{\mathrm{xc}}(n(\boldsymbol{r},t))}{\delta n(\boldsymbol{r},t)} + v_{\mathrm{ext}}(\boldsymbol{r},t)}_{v_{\mathrm{KS}}(\boldsymbol{r},t)}\Big]\phi_i(\boldsymbol{r},t).$$

Now KS operator depends on time t:

$$\textit{electronic density:} n(\boldsymbol{r},t) = \sum_{i=1}^{N_v}\phi_i(\boldsymbol{r},t)\phi_i^*(\boldsymbol{r},t).$$

G. Onida, L. Reining and A. Rubio, *Electronic excitations: density-functional versus many-body Green's function approaches*, Rev. Mod. Phys. 74, 2002, (59 pages).

Weak External Perturbation

$$\text{DFT:}\quad H_{\text{KS}}^{\text{GS}}\phi_i(\boldsymbol{r}) \equiv \left[-\frac{1}{2}\nabla^2 + v_{\text{KS}}(\boldsymbol{r})\right]\phi_i(\boldsymbol{r}) = \lambda_i\phi_i(\boldsymbol{r}),$$

$$v_{\text{KS}}(\boldsymbol{r}) = v_{\text{H}}(\boldsymbol{r}) + v_{\text{xc}}(\boldsymbol{r}) + v_{\text{ext}}(\boldsymbol{r}).$$

Perturb $v_{\text{ext}}(\boldsymbol{r})$ slightly to $v_{\text{ext}}(\boldsymbol{r}, t) = v_{\text{ext}}(\boldsymbol{r}) + \mathring{v}_{\text{ext}}(\boldsymbol{r}, t)$, which in turn induce perturbations to $v_{\text{Hxc}}(\boldsymbol{r}) \equiv v_{\text{H}}(\boldsymbol{r}) + v_{\text{xc}}(\boldsymbol{r})$:

$$v_{\text{Hxc}}(\boldsymbol{r}, t) = v_{\text{Hxc}}(\boldsymbol{r}) + \mathring{v}_{\text{Hxc}}(\boldsymbol{r}, t).$$

$$\text{TD-DFT:}\quad \iota\frac{\partial}{\partial t}\phi_i(\boldsymbol{r}, t) = H_{\text{KS}}(t)\phi_i(\boldsymbol{r}, t) \equiv \left[-\frac{1}{2}\nabla^2 + v_{\text{KS}}(\boldsymbol{r}, t)\right]\phi_i(\boldsymbol{r}, t),$$

$$\begin{aligned} H_{\text{KS}}(t) &= -\frac{1}{2}\nabla^2 + v_{\text{H}}(\boldsymbol{r}, t) + v_{\text{xc}}(\boldsymbol{r}, t) + v_{\text{ext}}(\boldsymbol{r}, t) \\ &= H_{\text{KS}}^{\text{GS}} + \mathring{v}_{\text{Hxc}}(\boldsymbol{r}, t) + \mathring{v}_{\text{ext}}(\boldsymbol{r}, t). \end{aligned}$$

Linear Response Theory

Seek information on first order change in $n(\boldsymbol{r},t) = n(\boldsymbol{r}) + \mathring{n}(\boldsymbol{r},t)$:

$$\begin{aligned}
\phi_i(\boldsymbol{r},t) &= \phi_i(\boldsymbol{r}) + \mathring{\phi}_i(\boldsymbol{r},t), \\
n(\boldsymbol{r},t) &\equiv n(\boldsymbol{r}) + \mathring{n}(\boldsymbol{r},t) \\
&= n(\boldsymbol{r}) + \sum_{i=1}^{N_v} \left[\mathring{\phi}_i^*(\boldsymbol{r},t)\phi_i(\boldsymbol{r}) + \phi_i^*(\boldsymbol{r})\mathring{\phi}_i(\boldsymbol{r},t) \right].
\end{aligned}$$

Better to explain using the single-particle density matrix which reads

$$\rho(\boldsymbol{r},t) = \sum_{i=1}^{N_v} |\phi_i(\boldsymbol{r},t)\rangle \langle \phi_i(\boldsymbol{r},t)| = \rho(\boldsymbol{r}) + \mathring{\rho}(\boldsymbol{r},t),$$

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N_v} |\phi_i(\boldsymbol{r})\rangle \langle \phi_i(\boldsymbol{r})|,$$

$$\mathring{\rho}(\boldsymbol{r},t) = \sum_{i=1}^{N_v} \left(\left|\mathring{\phi}_i(\boldsymbol{r},t)\right\rangle \langle \phi_i(\boldsymbol{r})| + |\phi_i(\boldsymbol{r})\rangle \left\langle \mathring{\phi}_i(\boldsymbol{r},t)\right| \right).$$

(For Dirac **Bra-ket** notation, google `bra-ket`.)

Linear Response Theory

Differentiate $\rho(\boldsymbol{r}, t)$ with respect to t to get

$$\begin{aligned}\iota\frac{\partial}{\partial t}\rho(\boldsymbol{r}, t) &= \sum_{i=1}^{N_v}\left[H_{\mathrm{KS}}(t)\left|\phi_i(\boldsymbol{r}, t)\right\rangle\left\langle\phi_i(\boldsymbol{r}, t)\right| - \left|\phi_i(\boldsymbol{r}, t)\right\rangle\left\langle\phi_i(\boldsymbol{r}, t)\right| H_{\mathrm{KS}}(t)\right] \\ &= \left[H_{\mathrm{KS}}(t), \rho(\boldsymbol{r}, t)\right].\end{aligned}$$

Substitute $H_{\mathrm{KS}}(t) = H_{\mathrm{KS}}^{\mathrm{GS}} + \mathring{v}_{\mathrm{Hxc}}(\boldsymbol{r}, t) + \mathring{v}_{\mathrm{ext}}(\boldsymbol{r}, t)$ and $\rho(\boldsymbol{r}, t) = \rho(\boldsymbol{r}) + \mathring{\rho}(\boldsymbol{r}, t)$ to get

$$\begin{aligned}\iota\frac{\partial}{\partial t}\mathring{\rho}(\boldsymbol{r}, t) &= \left[H_{\mathrm{KS}}^{\mathrm{GS}}, \mathring{\rho}(\boldsymbol{r}, t)\right] + \left[\mathring{v}_{\mathrm{Hxc}}(\boldsymbol{r}, t), \rho(\boldsymbol{r})\right] + \left[\mathring{v}_{\mathrm{ext}}(\boldsymbol{r}, t), \rho(\boldsymbol{r})\right] \\ &= \mathcal{L}\mathring{\rho}(\boldsymbol{r}, t) + \left[\mathring{v}_{\mathrm{ext}}(\boldsymbol{r}, t), \rho(\boldsymbol{r})\right],\end{aligned}$$

where $\mathcal{L}$ is the *Liouvillian super-operator*:

$$\mathcal{L}\mathring{\rho}(\boldsymbol{r}, t) := \left[H_{\mathrm{KS}}^{\mathrm{GS}}, \mathring{\rho}(\boldsymbol{r}, t)\right] + \left[\mathring{v}_{\mathrm{Hxc}}(\boldsymbol{r}, t), \rho(\boldsymbol{r})\right].$$

Linear Response Theory

$$\begin{aligned}\iota\frac{\partial}{\partial t}\mathring{\rho}(\boldsymbol{r},t) &= \left[H_{\mathrm{KS}}^{\mathrm{GS}},\mathring{\rho}(\boldsymbol{r},t)\right] + [\mathring{v}_{\mathrm{Hxc}}(\boldsymbol{r},t),\rho(\boldsymbol{r})] + [\mathring{v}_{\mathrm{ext}}(\boldsymbol{r},t),\rho(\boldsymbol{r})] \\ &= \mathcal{L}\mathring{\rho}(\boldsymbol{r},t) + [\mathring{v}_{\mathrm{ext}}(\boldsymbol{r},t),\rho(\boldsymbol{r})]\,,\end{aligned}$$

Apply the Fourier transformation to get

$$\begin{aligned}\omega\mathring{\rho}(\boldsymbol{r},\omega) &= \left[H_{\mathrm{KS}}^{\mathrm{GS}},\mathring{\rho}(\boldsymbol{r},\omega)\right] + [\mathring{v}_{\mathrm{Hxc}}(\boldsymbol{r},\omega),\rho(\boldsymbol{r})] + [\mathring{v}_{\mathrm{ext}}(\boldsymbol{r},\omega),\rho(\boldsymbol{r})]\,, \\ &= \mathcal{L}\mathring{\rho}(\boldsymbol{r},\omega) + [\mathring{v}_{\mathrm{ext}}(\boldsymbol{r},\omega),\rho(\boldsymbol{r})]\,,\end{aligned}$$

where $\mathcal{L}\mathring{\rho}(\boldsymbol{r},\omega) = \left[H_{\mathrm{KS}}^{\mathrm{GS}},\mathring{\rho}(\boldsymbol{r},\omega)\right] + [\mathring{v}_{\mathrm{Hxc}}(\boldsymbol{r},\omega),\rho(\boldsymbol{r})]$. Therefore

$$(\omega-\mathcal{L})\mathring{\rho}(\boldsymbol{r},\omega) = [\mathring{v}_{\mathrm{ext}}(\boldsymbol{r},\omega),\rho(\boldsymbol{r})]\,.$$

Set $\mathring{v}_{\mathrm{ext}}(\boldsymbol{r},\omega) = 0 \Rightarrow$ an eigenvalue problem; the smallest positive eigenvalues and associated eigenvectors give excitation states.

Matrix representation of $\mathcal{L}$

Can show:

$$\mathring{\rho}(\boldsymbol{r},t) = \begin{array}{c} \\ \mathcal{R}(\phi_1(r)) \\ \vdots \\ \mathcal{R}(\phi_{N_v}(r)) \\ \mathcal{N}(\rho(\boldsymbol{r})) \end{array} \begin{array}{c} \begin{array}{cccc} \mathcal{R}(\phi_1(r)) & \cdots & \mathcal{R}(\phi_{N_v}(r)) & \mathcal{N}(\rho(\boldsymbol{r})) \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 & \left\langle \mathring{\phi}_1(\boldsymbol{r},t) \right| \rho^{\perp}(\boldsymbol{r}) \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \left\langle \mathring{\phi}_{N_v}(\boldsymbol{r},t) \right| \rho^{\perp}(\boldsymbol{r}) \\ \rho^{\perp}(\boldsymbol{r}) \left| \mathring{\phi}_1(\boldsymbol{r},t) \right\rangle & \cdots & \rho^{\perp}(\boldsymbol{r}) \left| \mathring{\phi}_{N_v}(\boldsymbol{r},t) \right\rangle & 0 \end{pmatrix} \end{array}.$$

Hence basis functions of the "vector space" of all possible $\mathring{\rho}(\boldsymbol{r},t)$

$$x_i(\boldsymbol{r},t) = \rho^{\perp}(\boldsymbol{r}) \left| \mathring{\phi}_i(\boldsymbol{r},t) \right\rangle, \quad y_i(\boldsymbol{r},t) = \left\langle \mathring{\phi}_i(\boldsymbol{r},t) \right| \rho^{\perp}(\boldsymbol{r}).$$

In the frequency space:

$$\mathring{\rho}(\boldsymbol{r},\omega) = \begin{array}{c} \\ \mathcal{R}(\phi_1(r)) \\ \vdots \\ \mathcal{R}(\phi_{N_v}(r)) \\ \mathcal{N}(\rho(\boldsymbol{r})) \end{array} \begin{array}{c} \begin{array}{cccc} \mathcal{R}(\phi_1(r)) & \cdots & \mathcal{R}(\phi_{N_v}(r)) & \mathcal{N}(\rho(\boldsymbol{r})) \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 & y_1(\boldsymbol{r},\omega) \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & y_{N_v}(\boldsymbol{r},\omega) \\ x_1(\boldsymbol{r},\omega) & \cdots & x_{N_v}(\boldsymbol{r},\omega) & 0 \end{pmatrix} \end{array}.$$

Matrix representation of $\mathcal{L}$

$$\mathcal{L}\begin{pmatrix} x_1(\boldsymbol{r},\omega) \\ \vdots \\ x_{N_v}(\boldsymbol{r},\omega) \\ y_1(\boldsymbol{r},\omega) \\ \vdots \\ y_{N_v}(\boldsymbol{r},\omega) \end{pmatrix} = \begin{pmatrix} \mathcal{D}+\mathcal{K} & \mathcal{K} \\ -\mathcal{K} & -\mathcal{D}-\mathcal{K} \end{pmatrix} \begin{pmatrix} x_1(\boldsymbol{r},\omega) \\ \vdots \\ x_{N_v}(\boldsymbol{r},\omega) \\ y_1(\boldsymbol{r},\omega) \\ \vdots \\ y_{N_v}(\boldsymbol{r},\omega) \end{pmatrix},$$

$$\mathcal{D} = \mathrm{diag}\left(\rho^{\perp}(\boldsymbol{r})H_{\mathrm{KS}}^{\mathrm{GS}}\rho^{\perp}(\boldsymbol{r}) - \epsilon_1, \ldots, \rho^{\perp}(\boldsymbol{r})H_{\mathrm{KS}}^{\mathrm{GS}}\rho^{\perp}(\boldsymbol{r}) - \epsilon_{N_v}\right),$$

$$\mathcal{K}\begin{pmatrix} z_1(\boldsymbol{r}) \\ \vdots \\ z_{N_v}(\boldsymbol{r}) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N_v} \rho^{\perp}(\boldsymbol{r}) \int \kappa(\boldsymbol{r},\boldsymbol{r}')\phi_i(\boldsymbol{r}')\, z_i(\boldsymbol{r}')\, d\boldsymbol{r}' \;\; |\phi_1(\boldsymbol{r})\rangle \\ \vdots \\ \sum_{i=1}^{N_v} \rho^{\perp}(\boldsymbol{r}) \int \kappa(\boldsymbol{r},\boldsymbol{r}')\phi_i(\boldsymbol{r}')\, z_i(\boldsymbol{r}')\, d\boldsymbol{r}' \;\; |\phi_{N_v}(\boldsymbol{r})\rangle \end{pmatrix}.$$

Linear Response Eigenvalue Problem

First several smallest positive eigenvalues and corresponding eigenvectors of

$$\mathscr{H}\begin{bmatrix} u \\ v \end{bmatrix} \equiv \begin{bmatrix} A & B \\ -B & -A \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix},$$

$$A^{\mathsf{T}} = A,\ B^{\mathsf{T}} = B \in \mathbb{R}^{n\times n}, \quad \begin{bmatrix} A & B \\ B & A \end{bmatrix} \succ 0.$$

Equivalently, $Hz = \lambda z$:

$$J = \frac{1}{\sqrt{2}}\begin{bmatrix} I_n & I_n \\ I_n & -I_n \end{bmatrix}, \quad J^{\mathsf{T}}J = J^2 = I_{2n},$$

$$J^{\mathsf{T}}\begin{bmatrix} A & B \\ -B & -A \end{bmatrix} J = \begin{bmatrix} 0 & A-B \\ A+B & 0 \end{bmatrix} =: \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} =: H.$$

$$K = A - B \succ 0, \quad M = A + B \succ 0.$$

H non-symmetric, but rich structure to take advantage of.

References

For Linear Response Eigenvalue Problem in general


G. Onida, L. Reining, and A. Rubio. Electronic excitations: density-functional versus many-body Green's function approaches. *Rev. Mod. Phys*, 74(2):601–659, 2002.


Dario Rocca. *Time-Dependent Density Functional Perturbation Theory: New algorithms with Applications to Molecular Spectra*. PhD thesis, The International School for Advanced Studies, Trieste,Italy, 2007.


D. J. Thouless. Vibrational states of nuclei in the random phase approximation. *Nuclear Physics*, 22(1):78–95, 1961.

Material in what follows on Linear Response Eigenvalue Problem largely taken from


Zhaojun Bai and Ren-Cang Li. Minimization principle for linear response eigenvalue problem, I: Theory. *SIAM J. Matrix Anal. Appl.*, 33(4):1075–1100, 2012.


Zhaojun Bai and Ren-Cang Li. Minimization principle for linear response eigenvalue problem, II: Computation. *SIAM J. Matrix Anal. Appl.*, 34(2):392–416, 2013.


D. Rocca, Z. Bai, R.-C. Li, and G. Galli. A block variational procedure for the iterative diagonalization of non-Hermitian random-phase approximation matrices. *J. Chem. Phys.*, 136:034111, 2012.

Linear Response Eigenvalue Problem

First several smallest positive eigenvalues, eigenvectors of

$$\mathscr{H}\begin{bmatrix} u \\ v \end{bmatrix} \equiv \begin{bmatrix} A & B \\ -B & -A \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix},$$

$A^{\mathsf{T}} = A$, $B^{\mathsf{T}} = B \in \mathbb{R}^{n\times n}$, $\begin{bmatrix} A & B \\ B & A \end{bmatrix}$ positive definite.

$$J = \frac{1}{\sqrt{2}}\begin{bmatrix} I_n & I_n \\ I_n & -I_n \end{bmatrix}, \quad J^{\mathsf{T}} J = J^2 = I_{2n},$$

$$J^{\mathsf{T}}\begin{bmatrix} A & B \\ -B & -A \end{bmatrix} J = \begin{bmatrix} 0 & A-B \\ A+B & 0 \end{bmatrix} =: \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} =: H.$$

$K = A - B$, $M = A + B \in \mathbb{R}^{n\times n}$ definite because

$$J^{\mathsf{T}}\begin{bmatrix} A & B \\ B & A \end{bmatrix} J = \begin{bmatrix} A+B & 0 \\ 0 & A-B \end{bmatrix} \equiv \begin{bmatrix} M & \\ & K \end{bmatrix}.$$

Equivalent Forms

Eigenvalue problem for $\mathscr{H}$ – original LR:

$$\mathscr{H}\begin{bmatrix} u \\ v \end{bmatrix} \equiv \begin{bmatrix} A & B \\ -B & -A \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix},$$

Eigenvalue problem for H – transformed LR:

$$Hz \equiv \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix}\begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} y \\ x \end{bmatrix} \equiv \lambda z,$$

Eigenvalue Problems for $\mathscr{H}$ and H equivalent:

- Same eigenvalues, and
- Eigenvectors related by

$$\begin{bmatrix} y \\ x \end{bmatrix} = J^{\mathsf{T}}\begin{bmatrix} u \\ v \end{bmatrix}, \quad \begin{bmatrix} u \\ v \end{bmatrix} = J\begin{bmatrix} y \\ x \end{bmatrix}$$

Problem. $H = \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix}$, $0 \prec K^{\mathsf{T}} = K, 0 \prec M^{\mathsf{T}} = M \in \mathbb{R}^{n \times n}$.

KM and MK have positive eigenvalues:

$$0 < \lambda_1^2 \leq \lambda_2^2 \leq \cdots \leq \lambda_n^2,$$

where $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, because

$$\text{eig}(KM) = \text{eig}(K^{1/2} K^{1/2} M) = \text{eig}(K^{1/2} M K^{1/2}).$$

$H^2 = \begin{bmatrix} KM & 0 \\ 0 & MK \end{bmatrix} \Rightarrow H$ has eigenvalues

$$-\lambda_n \leq \cdots \leq -\lambda_1 < +\lambda_1 \leq \cdots \leq +\lambda_n.$$

Eigendecomposition

$\exists\, X = Y^{-\mathsf{T}} \in \mathbb{R}^{n\times n}$ such that

$$K = Y\Lambda^2 Y^{\mathsf{T}}, \quad M = XX^{\mathsf{T}}, \quad \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Critically important later.

Proof.

1) Cholesky decomposition: $M^{-1} = R^{\mathsf{T}} R$.
2) Eigendecomposition: $R^{-\mathsf{T}} K R^{-1} = Q\Lambda^2 Q^{\mathsf{T}}$, $Q^{\mathsf{T}} Q = I_n$.
3) Finally $Y = R^{\mathsf{T}} Q$, $X = Y^{-\mathsf{T}}$. □

H is diagonalizable with Eigendecomposition:

$$H \begin{bmatrix} Y\Lambda & Y\Lambda \\ X & -X \end{bmatrix} = \begin{bmatrix} Y\Lambda & Y\Lambda \\ X & -X \end{bmatrix} \begin{bmatrix} \Lambda & \\ & -\Lambda \end{bmatrix}.$$

Thouless' Minimization Principle

Eigenvalue problem for $\mathscr{H}$: special case of *Hamiltonian eigenvalue problem*.

Eigenvalues appear in $\pm\lambda$ pairs:

$$-\lambda_n \leq \cdots \leq -\lambda_1 < +\lambda_1 \leq \cdots \leq +\lambda_n.$$

Thouless' Minimization Principle (1961):

$$\lambda_1 = \min_{u,v} \varrho(u,v), \quad \varrho(u,v) = \frac{\begin{bmatrix} u \\ v \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}}{|u^{\mathsf{T}}u - v^{\mathsf{T}}v|}.$$

Many of today's minimization approaches for computing λ_1 are results of this principle.

Thouless' Minimization Principle

Use $\begin{bmatrix} y \\ x \end{bmatrix} = J^{\mathsf{T}} \begin{bmatrix} u \\ v \end{bmatrix}$, $\begin{bmatrix} u \\ v \end{bmatrix} = J \begin{bmatrix} y \\ x \end{bmatrix}$ to get

Thouless' Minimization Principle (in different form)

$$\lambda_1 = \min_{x,y} \rho(x,y), \quad \rho(x,y) \stackrel{\text{def}}{=} \frac{x^{\mathsf{T}} K x + y^{\mathsf{T}} M y}{2|x^{\mathsf{T}} y|}.$$

Will call both $\varrho(u,v)$ and $\rho(x,y)$ *Thouless' Functional*.

Proof.

Recall $K = Y \Lambda^2 Y^{\mathsf{T}}$, $M = XX^{\mathsf{T}}$, and $X = Y^{-\mathsf{T}}$. We have

$$\begin{aligned} \min_{x,y} \frac{x^{\mathsf{T}} K x + y^{\mathsf{T}} M y}{2|x^{\mathsf{T}} y|} &= \min_{x,y} \frac{x^{\mathsf{T}} Y \Lambda^2 Y^{\mathsf{T}} x + y^{\mathsf{T}} Y^{-\mathsf{T}} Y^{-1} y}{2|x^{\mathsf{T}} Y Y^{-1} y|} \\ &= \min_{\widetilde{x},\widetilde{y}} \frac{\widetilde{x}^{\mathsf{T}} \Lambda^2 \widetilde{x} + \widetilde{y}^{\mathsf{T}} \widetilde{y}}{2|\widetilde{x}^{\mathsf{T}} \widetilde{y}|} \\ &\geq \min_{\widetilde{x},\widetilde{y}} \frac{2\sum_i \lambda_i |\widetilde{x}_{(i)} \widetilde{y}_{(i)}|}{2|\sum_i \widetilde{x}_{(i)} \widetilde{y}_{(i)}|} \geq \lambda_1. \end{aligned}$$

Careful analysis $\Rightarrow$ equality signs realizable, and optimal argument pair produces eigenvector. □

Previous Work - summary

Four decades' researches by computational (quantum) physicists and chemists and numerical analysts.

Following three eigenvalue problems are equivalent:

$$Hz \equiv \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} y \\ x \end{bmatrix} \equiv \lambda z, \tag{Eig-H}$$

$$KMy = \lambda^2 y, \tag{Eig-KM}$$

$$MKx = \lambda^2 x. \tag{Eig-MK}$$

By computational (quantum) physicists and chemists:

- Chi (1970): solve (Eig-KM) through Symmetric Eigenvalue Problem (SEP) RKR^{T}, where $M = R^{\mathsf{T}}R$ (Cholesky decomposition).
- Davidson-type algorithms (1980s & 1990s), Lanczos-like algorithms (1990s & 2000s)
- CG-like algorithms (more recently, based on Thouless' principle)

By numerical analysts:

- Wilkinson (1960s) discussed (Eig-KM) and (Eig-MK). Implemented as LAPACK `xSYGVD`
- GR algorithm for product eigenvalue problems, generalizing well-known QR algorithm (Watkins, Kressner)
- Krylov-Schur, Jacobi-Davidson, Hamiltonian Krylov-Schur-type, symplectic Lanczos, ...

Previous Work & What to Come

Trend. Huge size – n in the order 10^6 or larger; pose tremendous challenge.

Despite four decades' researches, it is still challenging to robustly and efficiently compute first several positive eigenvalues and eigenvectors.

To come:

- New theory for H that parallels Symmetric Eigenvalue Problem (SEP)
- New algorithms capable of computing first several positive eigenvalues and eigenvectors *simultaneously*.

Deflating Subspaces

$\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^n$, subspaces. Call $\{\mathcal{U}, \mathcal{V}\}$ a *pair of deflating subspaces* of $\{K, M\}$ if

$$K\mathcal{U} \subseteq \mathcal{V} \quad \text{and} \quad M\mathcal{V} \subseteq \mathcal{U}.$$

Let $U \in \mathbb{R}^{n\times k}$, $V \in \mathbb{R}^{n\times k}$, basis matrices for $\mathcal{U}$ and $\mathcal{V}$, resp.

$\exists\ K_R, M_R \in \mathbb{R}^{k\times k}$ such that

$$KU = VK_R, \quad MV = UM_R.$$

In fact, for left generalized inverses $U^\dashv$, $V^\dashv$ of U, V, resp.,

$$K_R = V^\dashv KU, \quad M_R = U^\dashv MV.$$

For example, $U^\dashv U = I$ for

$$U^\dashv = (U^T U)^{-1} U^T, \quad \text{but we prefer}$$
$$U^\dashv = (V^T U)^{-1} V^T \quad \text{if } (V^T U)^{-1} \text{ exists.}$$

Basics: Deflating Subspaces

$$KU = VK_{\mathrm{R}},\ MV = UM_{\mathrm{R}} \Rightarrow \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} \begin{bmatrix} V & \\ & U \end{bmatrix} = \begin{bmatrix} V & \\ & U \end{bmatrix} \begin{bmatrix} 0 & K_{\mathrm{R}} \\ M_{\mathrm{R}} & 0 \end{bmatrix}$$

$H_{\mathrm{R}} = \begin{bmatrix} 0 & K_{\mathrm{R}} \\ M_{\mathrm{R}} & 0 \end{bmatrix}$ is a restriction of H onto $\mathcal{V} \oplus \mathcal{U}$.

H_{R} same block structure as H; but lose symmetry in K, M.

Suppose $W \stackrel{\mathrm{def}}{=} U^{\mathsf{T}}V$ nonsingular. Factorize $W = W_1^{\mathsf{T}} W_2$, where W_i nonsingular. Define

$$H_{\mathrm{SR}} = \begin{bmatrix} 0 & W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \\ W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} & 0 \end{bmatrix},$$

another restriction of H onto $\mathcal{V} \oplus \mathcal{U}$, too:

$$H \begin{bmatrix} VW_2^{-1} & \\ & UW_1^{-1} \end{bmatrix} = \begin{bmatrix} VW_2^{-1} & \\ & UW_1^{-1} \end{bmatrix} H_{\mathrm{SR}}.$$

H_{SR} same block structure as H and retain symmetry in K, M. Major role to come.

Trace Minimization Principle

Trace Minimization Principle

$$\sum_{i=1}^{k} \lambda_i = \frac{1}{2} \min_{U^T V = I_k} \operatorname{trace}(U^T K U + V^T M V).$$

If $\lambda_k < \lambda_{k+1}$, optimal $\{\operatorname{span}(U), \operatorname{span}(V)\}$ gives deflating subspaces of $\{K, M\}$ corresponding to $\pm\lambda_i$, $1 \le i \le k$.

Quite similar to the Trace Minimization Principle for Symmetric Eigenvalue Problem (SEP) discussed earlier.

A lengthy proof can be found in


Zhaojun Bai and Ren-Cang Li. Minimization principle for linear response eigenvalue problem, I: Theory. *SIAM J. Matrix Anal. Appl.*, 33(4):1075–1100, 2012.

Cauchy-like Interlacing Inequalities

Cauchy-like Interlacing Inequalities

$U, V \in \mathbb{R}^{n\times k}$; $W = U^{\mathsf{T}}V$ nonsingular; $W = W_1^{\mathsf{T}} W_2$, $W_i \in \mathbb{R}^{k\times k}$ nonsingular; $\mathcal{U} = \text{span}(U)$, $\mathcal{V} = \text{span}(V)$; Eigenvalues of

$$H_{\text{SR}} = \begin{bmatrix} 0 & W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \\ W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} & 0 \end{bmatrix}.$$

are $\pm\mu_i$ $(1 \le i \le k)$, where $0 \le \mu_1 \le \cdots \le \mu_k$. Then for $1 \le i \le k$

$$\lambda_i \le \mu_i \le \min\left\{ \lambda_{i+2(n-k)}, \frac{\sqrt{\min\{\kappa(K), \kappa(M)\}}}{\cos\angle(\mathcal{U}, \mathcal{V})} \lambda_{i+n-k} \right\},$$

where $\lambda_j = \infty$ for $j > n$. If either $\mathcal{U} = M\mathcal{V}$ or $\mathcal{V} = K\mathcal{U}$, then $\lambda_i \le \mu_i \le \lambda_{i+n-k}$.

Quite similar to the Cauchy Interlacing Inequalities for Symmetric Eigenvalue Problem (SEP) discussed earlier.

A lengthy proof can be found in Bai and Li (2012) mentioned in the previous slide.

Best Approximation

Seeking "*best possible*" approximations from the suitably built subspaces.

Given $\{\mathcal{U}, \mathcal{V}\}$, a pair of subspaces, $\dim(\mathcal{U}) = \dim(\mathcal{V}) = n_b$.

Minimization principles motivate us to seek

- the best approximation to λ_1 in the sense of

$$\min_{x \in \mathcal{U},\, y \in \mathcal{V}} \rho(x, y)$$

  and its associated approximate eigenvector;

- the best approximations to λ_j $(1 \leq j \leq k)$ in the sense of

$$\frac{1}{2} \min_{\substack{\mathrm{span}(\widehat{U}) \subseteq \mathcal{U},\, \mathrm{span}(\widehat{V}) \subseteq \mathcal{V} \\ \widehat{U}^{\mathsf{T}} \widehat{V} = I_k}} \mathrm{trace}(\widehat{U}^{\mathsf{T}} K \widehat{U} + \widehat{V}^{\mathsf{T}} M \widehat{V})$$

  and their associated approximate eigenvectors. Necessarily $k \leq n_b$.

Best Approximation: λ_1

U, $V \in \mathbb{R}^{n \times n_b}$, basis matrices of $\mathcal{U}$ and $\mathcal{V}$. Assume $W = U^{\mathsf{T}} V$ nonsingular. Factorize $W = W_1^{\mathsf{T}} W_2$, $W_i \in \mathbb{R}^{n_b \times n_b}$ nonsingular.

$x \in \mathcal{U}$, $y \in \mathcal{V} \Leftrightarrow x = U\hat{u}$, $y = V\hat{v}$ for $\hat{u}$, $\hat{v} \in \mathbb{R}^{n_b}$.

$$\begin{aligned} \rho(x,y) &= \frac{x^{\mathsf{T}} K x + y^{\mathsf{T}} M y}{2|x^{\mathsf{T}} y|} = \frac{\hat{u}^{\mathsf{T}} U^{\mathsf{T}} K U \hat{u} + \hat{v}^{\mathsf{T}} V^{\mathsf{T}} M V \hat{v}}{2|\hat{u}^{\mathsf{T}} W \hat{v}|} \\ &= \frac{\hat{x}^{\mathsf{T}} W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \hat{x} + \hat{y}^{\mathsf{T}} W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} \hat{y}}{2|\hat{x}^{\mathsf{T}} \hat{y}|}, \end{aligned}$$

where $\hat{x} = W_1 \hat{u}$ and $\hat{y} = W_2 \hat{v}$.

$$\min_{x \in \mathcal{U},\, y \in \mathcal{V}} \rho(x,y) = \min_{\hat{x},\hat{y}} \frac{\hat{x}^{\mathsf{T}} W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \hat{x} + \hat{y}^{\mathsf{T}} W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} \hat{y}}{2|\hat{x}^{\mathsf{T}} \hat{y}|}$$

which is the smallest positive eigenvalue of

$$H_{\mathrm{SR}} = \begin{bmatrix} 0 & W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \\ W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} & 0 \end{bmatrix}.$$

Best Approximation: λ_j $(1 \le j \le k)$

$$\left\{ \begin{array}{c} \mathrm{span}(\widehat{U}) \subseteq \mathcal{U},\ \mathrm{span}(\widehat{V}) \subseteq \mathcal{V}, \\ \widehat{U}^{\mathsf{T}} \widehat{V} = I_k \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \widehat{U} = U W_1^{-1} \widehat{X},\ \widehat{V} = V W_2^{-1} \widehat{Y}, \\ \widehat{X},\ \widehat{Y} \in \mathbb{R}^{n_b \times k},\ \widehat{X}^{\mathsf{T}} \widehat{Y} = I_k \end{array} \right\}$$

$$\widehat{U}^{\mathsf{T}} K \widehat{U} + \widehat{V}^{\mathsf{T}} M \widehat{V} = \widehat{X}^{\mathsf{T}} W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \widehat{X} + \widehat{Y}^{\mathsf{T}} W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} \widehat{Y}.$$

$$\begin{aligned} &\min_{\substack{\mathrm{span}(\widehat{U}) \subseteq \mathcal{U},\, \mathrm{span}(\widehat{V}) \subseteq \mathcal{V} \\ \widehat{U}^{\mathsf{T}} \widehat{V} = I_k}} \mathrm{trace}(\widehat{U}^{\mathsf{T}} K \widehat{U} + \widehat{V}^{\mathsf{T}} M \widehat{V}) \\ &\quad = \min_{\widehat{X}^{\mathsf{T}} \widehat{Y} = I_k} \mathrm{trace}(\widehat{X}^{\mathsf{T}} W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \widehat{X} + \widehat{Y}^{\mathsf{T}} W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} \widehat{Y}). \end{aligned}$$

which is the sum of 1st k smallest positive eigenvalue of

$$H_{\mathrm{SR}} = \begin{bmatrix} 0 & W_1^{-\mathsf{T}} U^{\mathsf{T}} K U W_1^{-1} \\ W_2^{-\mathsf{T}} V^{\mathsf{T}} M V W_2^{-1} & 0 \end{bmatrix}.$$

Best Approximation: Eigenvectors

Positive eigenvalues of H_{SR}: $0 \le \rho_1 \le \cdots \le \rho_{n_b}$. Associated eigenvectors $\hat{z}_j$.

$$H_{\mathrm{SR}}\hat{z}_j = \rho_j \hat{z}_j, \quad \hat{z}_j = \begin{bmatrix} \hat{y}_j \\ \hat{x}_j \end{bmatrix}.$$

Then $\rho(UW_1^{-1}\hat{x}_j, VW_2^{-1}\hat{y}_j) = \rho_j$ for $j = 1, \ldots, n_b$.

Naturally, take $\lambda_j \approx \rho_j$, and corresponding approximate eigenvectors of H:

$$\widetilde{z}_j \equiv \begin{bmatrix} \widetilde{y}_j \\ \widetilde{x}_j \end{bmatrix} = \begin{bmatrix} VW_2^{-1}\hat{y}_j \\ UW_1^{-1}\hat{x}_j \end{bmatrix} \quad \text{for } j = 1, \ldots, n_b.$$

What if $U^{\mathsf{T}}V$ is singular? Still can do, just more complicated

Zhaojun Bai and Ren-Cang Li. Minimization principle for linear response eigenvalue problem, II: Computation. *SIAM J. Matrix Anal. Appl.*, 34(2):392–416, 2013.

Gradients

Perturb x, y to $\hat{x} = x + p$, $\hat{y} = y + q$, p and q tiny . Assume $x^{\mathsf{T}} y \neq 0$.

Up to the first order in p and q,

$$\begin{aligned}
\rho(\hat{x}, \hat{y}) &= \frac{(x+p)^{\mathsf{T}} K (x+p) + (y+q)^{\mathsf{T}} M (y+q)}{2|(x+p)^{\mathsf{T}}(y+q)|} \\
&= \frac{x^{\mathsf{T}} K x + 2p^{\mathsf{T}} K x + y^{\mathsf{T}} M y + 2q^{\mathsf{T}} M y}{2|x^{\mathsf{T}} y + p^{\mathsf{T}} y + q^{\mathsf{T}} x|} \\
&= \frac{x^{\mathsf{T}} K x + 2p^{\mathsf{T}} K x + y^{\mathsf{T}} M y + 2q^{\mathsf{T}} M y}{2\,|x^{\mathsf{T}} y|} \left[1 - \frac{p^{\mathsf{T}} y + q^{\mathsf{T}} x}{x^{\mathsf{T}} y}\right] \\
&= \rho(x,y) + \frac{1}{x^{\mathsf{T}} y}\, p^{\mathsf{T}} \left[Kx - \rho(x,y)\, y\right] + \frac{1}{x^{\mathsf{T}} y}\, q^{\mathsf{T}} \left[My - \rho(x,y)\, x\right].
\end{aligned}$$

Partial gradients: $\nabla_x \rho = \dfrac{1}{x^{\mathsf{T}} y} \left[Kx - \rho(x,y)\, y\right]$, $\nabla_y \rho = \dfrac{1}{x^{\mathsf{T}} y} \left[My - \rho(x,y)\, x\right]$.

Closely related to residual:

$$Hz - \rho(x,y) z \equiv \begin{bmatrix} 0 & K \\ M & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} - \rho(x,y) \begin{bmatrix} y \\ x \end{bmatrix} = x^{\mathsf{T}} y \begin{bmatrix} \nabla_x \rho \\ \nabla_y \rho \end{bmatrix}.$$

4-D Search

Interested in solving $\min_{x,y} \rho(x,y) = \min_{x,y} \dfrac{x^T K x + y^T M y}{2|x^T y|}$ to compute λ_1.

Standard line search: Given current position $\begin{bmatrix} y \\ x \end{bmatrix}$, search direction $\begin{bmatrix} q \\ p \end{bmatrix}$, seek to minimize ρ along line

$$\left\{ \begin{bmatrix} y \\ x \end{bmatrix} + t \begin{bmatrix} q \\ p \end{bmatrix} \ : \ t \in \mathbb{R} \right\}$$

Doable via Calculus. But not flexible enough to have subspace extensions.

We will do differently.

Minimize ρ within the *4-dimensional subspace*

$$\left\{ \begin{bmatrix} \beta y + tq \\ \alpha x + sp \end{bmatrix} \quad \text{for all scalars } \alpha,\ \beta,\ s, \text{ and } t \right\}$$

to get

$$\min_{\alpha,\beta,s,t} \rho(\alpha x + sp, \beta y + tq) = \min_{u \in \text{span}(U),\ v \in \text{span}(V)} \rho(u,v),$$

where $U = [x, p]$ and $V = [y, q]$. Returned to *Best Approximation*.

4-D Search

Naturally take

$$\begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} \nabla_y \rho \\ \nabla_x \rho \end{bmatrix},$$

as in the *standard steepest descent* (SD) algorithm.

- Lead to plain 4-D SD algorithm for H
- Can design block versions for computing several eigenpairs
- Can incorporate pre-conditioners

All can be viewed as variants of locally optimal 4-D CG algorithms which we will discuss.

Locally Optimal 4-D CG

Notation: ℓ iteration index; j eigenpair index.
Standard: search next approximations within

$$\text{span}\left\{ \begin{bmatrix} y_j^{(\ell)} \\ x_j^{(\ell)} \end{bmatrix}, \begin{bmatrix} y_j^{(\ell-1)} \\ x_j^{(\ell-1)} \end{bmatrix}, \begin{bmatrix} q_j \\ p_j \end{bmatrix}, \quad j = 1:k \right\},$$

where

$$\begin{bmatrix} q_j \\ p_j \end{bmatrix} = \Phi \begin{bmatrix} \nabla_x \rho \\ \nabla_y \rho \end{bmatrix} \Big|_{(x,y)=(x_j^{(\ell)}, y_j^{(\ell)})},$$

and Φ is a preconditioner to be discussed later.

We do differently: search next approximations within

$$\text{span}\left\{ \begin{bmatrix} y_j^{(\ell)} \\ 0 \end{bmatrix}, \begin{bmatrix} y_j^{(\ell-1)} \\ 0 \end{bmatrix}, \begin{bmatrix} q_j \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x_j^{(\ell)} \end{bmatrix}, \begin{bmatrix} 0 \\ x_j^{(\ell-1)} \end{bmatrix}, \begin{bmatrix} 0 \\ p_j \end{bmatrix} \quad j = 1:k \right\}.$$

Breaking vectors into two this way is a common technique today in developing structure-preserving alg.:


Kevin J. Kerns and Andrew T. Yang. Preservation of passivity during RLC network reduction via split congruence transformations. *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, 17(7):582–591, July 1998.


R. W. Freund. SPRIM: Structure-preserving reduced-order interconnect macromodeling. In *Proc. Int. Conf. Computer Aided Design*, pages 80–87, Nov. 2004.


Ren-Cang Li and Zhaojun Bai. Structure-preserving model reductions using a Krylov subspace projection formulation. *Commun. Math. Sciences*, 3(2):179–199, 2005.

Locally Optimal Block Preconditioned 4-D CG

Locally Optimal Block Preconditioned 4-D CG

Given an initial approximation $Z_0 = [X_0^T, Y_0^T]^T \in \mathbb{C}^{2n \times n_b}$ with $\text{rank}(X_0) = \text{rank}(Y_0) = n_b$, the algorithm attempts to compute approximate eigenpair to (λ_j, z_j) for $1 \le j \le n_b$.

1: If jth column of Z_0 isn't an approximation to z_j already, compute initial approximation with $\{\mathcal{U}, \mathcal{V}\} := \{\mathcal{R}(X_0), \mathcal{R}(Y_0)\}$ to give a new Z_0;
2: **for** $\ell = 0, 1, \ldots$ **do**
3: test convergence and lock up the converged (to discuss later)
4: construct a preconditioner Φ_ℓ;
5: $\begin{bmatrix} Q_\ell \\ P_\ell \end{bmatrix} \leftarrow \Phi_\ell \begin{bmatrix} KX_\ell - Y_\ell \, \Omega_\ell \\ MY_\ell - X_\ell \, \Omega_\ell) \end{bmatrix}$, where $\Omega_\ell = \text{diag}(\rho_{\ell;j})$;
6: $U = (X_\ell, X_{\ell-1}, P_\ell)$, $V = (Y_\ell, Y_{\ell-1}, Q_\ell)$ (drop $X_{\ell-1}$ and $Y_{\ell-1}$ for $\ell = 0$);
7: orthogonalize the columns of U and V and decompose $W = U^{\mathsf{T}} V = W_1^{\mathsf{T}} W_2$;
8: construct H_{SR} (assume W is nonsingular);
9: compute n_b smallest positive eigenvalues $\rho_{\ell+1;j}$ of H_{SR}, and associated eigenvectors $\hat{z}_j$;
10: $Z_{\ell+1} := \begin{bmatrix} Y_{\ell+1} \\ X_{\ell+1} \end{bmatrix} = \begin{bmatrix} VW_2^{-1}[\hat{y}_1, \ldots, \hat{y}_k] \\ UW_1^{-1}[\hat{x}_1, \ldots, \hat{x}_k] \end{bmatrix}$ (normalize each column).
11: **end for**
12: **return** approximate eigenpairs to (λ_j, z_j) for $1 \le j \le n_b$.

Generic Pre-conditioner

For convenience, drop iteration index ℓ.

To compute eigenvalues close to μ: $\Phi = (H - \mu I_{2n})^{-1}$, and

$$\begin{bmatrix} Q \\ P \end{bmatrix} = \Phi R, \quad R = HZ - Z\Omega = \begin{bmatrix} KX - Y\,\Omega \\ MY - X\,\Omega \end{bmatrix},$$

one step of the inverse power iteration on the residual.

Interested in the smallest positive eigenvalues, naturally $\mu = 0$:

$$\Phi R = \begin{bmatrix} 0 & M^{-1} \\ K^{-1} & 0 \end{bmatrix} R = \begin{bmatrix} M^{-1}[MY - X\ \mathrm{diag}(\rho_j)] \\ K^{-1}[KX - Y\ \mathrm{diag}(\rho_j)] \end{bmatrix}.$$

Both P and Q computable (column-by-column) by (linear) CG.

In general for $\mu \neq 0$, multiplying by Φ involves solving linear system: $(H - \mu I_{2n})z = b$. Next slides consider this for $0 < \mu < \lambda_1$.

Generic Pre-conditioner: $(H - \mu I_{2n})z = b$

Can verify

$$\begin{bmatrix} I & 0 \\ M & \mu I \end{bmatrix} (H - \mu I) = \begin{bmatrix} I & 0 \\ M & \mu I \end{bmatrix} \begin{bmatrix} -\mu I & K \\ M & -\mu I \end{bmatrix} = \begin{bmatrix} -\mu I & K \\ 0 & MK - \mu^2 I \end{bmatrix}$$

to get

$$(H - \mu I_{2n})^{-1} = \begin{bmatrix} -\mu I & K \\ 0 & MK - \mu^2 I \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ M & \mu I \end{bmatrix}$$

Write $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ and $z = \begin{bmatrix} y \\ x \end{bmatrix}$. $(H - \mu I_{2n})z = b$ can be solved as

> solve $(MK - \mu^2 I)x = Mb_1 - \mu b_2$ for x, and then $y = \frac{1}{-\mu}(b_1 - Kx)$.

Remain to solve $(MK - \mu^2 I)x = c$ efficiently. Write $A = MK - \mu^2 I$, and symbolically

$$A = M^{1/2} \underbrace{(M^{1/2} K M^{1/2} - \mu^2 I)}_{=:\widehat{A}} M^{-1/2} = M^{1/2} \widehat{A} M^{-1/2}$$

$Ax = c$ is equivalent to $M^{1/2}\widehat{A}M^{-1/2}x = c \quad \Leftrightarrow \quad \widehat{A}\underbrace{M^{-1/2}x}_{=:\hat{x}} = \underbrace{M^{-1/2}c}_{=:\hat{c}}$.

$\widehat{A}$ is SPD because $0 < \mu < \lambda_1$. Can apply linear CG to $\widehat{A}\hat{x} = \hat{c}$ symbolically first and then translate to $Ax = c$.

$Ax \equiv (MK - \mu^2 I)x = c$

Transform to $\widehat{A}\hat{x} = \hat{c}$, $\widehat{A} = M^{-1/2} A M^{1/2}$, $\hat{x} = M^{-1/2} x$, $\hat{c} = M^{-1/2} x$.

Linear CG to $\widehat{A}\hat{x} = \hat{c}$: $\hat{r}_0 = \widehat{A}\hat{x}_0 - \hat{c}$, $\hat{p}_0 = -\hat{r}_0$, and for $i \geq 0$

$$\hat{x}_{i+1} = \hat{x}_i + \alpha_i \hat{p}_i, \;\; \hat{r}_{i+1} = \hat{r}_i + \alpha_i \widehat{A}\hat{p}_i, \;\; \hat{p}_{i+1} = -\hat{r}_{i+1} + \beta_i \hat{p}_i$$

where $\alpha_i = -\dfrac{\hat{p}_i^T \hat{r}_i}{\hat{p}_i^T \widehat{A} \hat{p}_i} = \dfrac{\hat{r}_i^T \hat{r}_i}{\hat{p}_i^T \widehat{A} \hat{p}_i}$, $\beta_i = \dfrac{\hat{p}_i^T \widehat{A} \hat{r}_{i+1}}{\hat{p}_i^T \widehat{A} \hat{p}_i} = \dfrac{\hat{r}_{i+1}^T \hat{r}_{i+1}}{\hat{r}_i^T \hat{r}_i}$.

To convert them back to x-space (so-to-speak): Note
$\hat{r} := \widehat{A}\hat{x} - \hat{c} = M^{-1/2}(Ax - c) =: M^{-1/2} r$. So $M^{1/2}\hat{p}_0 = -r_0$, and for $i \geq 0$,

$$x_{i+1} = x_i + \alpha_i M^{1/2} \hat{p}_i, \quad r_{i+1} = r_i + \alpha_i A M^{1/2} \hat{p}_i, \quad M^{1/2} \hat{p}_{i+1} = -r_{i+1} + \beta_i M^{1/2} \hat{p}_i.$$

Two possible choices for p-vectors (drop subscripts):

$$p = M^{1/2} \hat{p} \qquad \text{(natural)},$$
$$p = M^{-1/2} \hat{p} \qquad \text{(not-so-natural)}.$$

Difference in new formulas for α_i and β_i.

CG(I) for $Ax \equiv (MK - \mu^2 I)x = c$

Take $p = M^{1/2}\hat{p}$ (natural). Already $\hat{r} = M^{-1/2}r$, $\widehat{A} = M^{-1/2}AM^{1/2}$. So

$$\hat{p}^T\hat{r} = p^T M^{-1} r, \quad \hat{p}^T\widehat{A}\hat{p} = p^T M^{-1} A p, \quad \hat{r}^T\hat{r} = r^T M^{-1} r, \quad \hat{p}^T\widehat{A}\hat{r} = p^T M^{-1} A r.$$

Therefore

$$\alpha_i = -\frac{p_i^T M^{-1} r_i}{p_i^T M^{-1} A p_i} = \frac{r_i^T M^{-1} r_i}{p_i^T M^{-1} A p_i}, \quad \beta_i = \frac{p_i^T M^{-1} A r_{i+1}}{p_i^T M^{-1} A p_i} = \frac{r_{i+1}^T M^{-1} r_{i+1}}{r_i^T M^{-1} r_i}.$$

CG(I) for $Ax \equiv (MK - \mu^2 I)x = c$

Given an initial approximation x_0, a relative tolerance `rtol`, the algorithm solves $Ax \equiv (MK - \mu^2 I)x = c$.

1: $r_0 = Ax_0 - c$, $p_0 = -r_0$;
2: **for** $i = 0, 1, \ldots$ **do**
3: $\quad q_i = M^{-1}p_i$ by (linear) CG;
4: $\quad \alpha_i = -(q_i^T r_i)/(q_i^T A p_i)$, $x_{i+1} = x_i + \alpha_i p_i$, $r_{i+1} = r_i + \alpha_i A p_i$;
5: $\quad$ if $\|r_{i+1}\|_1/\|c\|_1 \le$ `rtol`, **BREAK**;
6: $\quad \beta_i = (q_i^T A r_{i+1})/(q_i^T A p_i)$, $p_{i+1} = -r_{i+1} + \beta_i p_i$;
7: **end for**
8: **return** last x_i as an approximate solution.

CG(II) for $Ax \equiv (MK - \mu^2 I)x = c$

Take $p = M^{-1/2}\hat{p}$ (not-so-natural). Already $\hat{r} = M^{-1/2}r$, $\widehat{A} = M^{-1/2}AM^{1/2}$. So

$$\hat{p}^T\hat{r} = p^T r, \quad \hat{p}^T\widehat{A}\hat{p} = p^T AMp, \quad \hat{r}^T\hat{r} = r^T M^{-1} r, \quad \hat{p}^T\widehat{A}\hat{r} = p^T Ar.$$

Therefore

$$\alpha_i = -\frac{p_i^T r_i}{p_i^T AMp_i} = \frac{r_i^T M^{-1} r_i}{p_i^T AMp_i}, \quad \beta_i = \frac{p_i^T Ar_{i+1}}{p_i^T AMp_i} = \frac{r_{i+1}^T M^{-1} r_{i+1}}{r_i^T M^{-1} r_i}. \tag{9}$$

CG(II) for $Ax \equiv (MK - \mu^2 I)x = c$

Given an initial approximation x_0, a relative tolerance `rtol`, the algorithm solves $Ax \equiv (MK - \mu^2 I)x = c$.

1: $r_0 = Ax_0 - c$, $q_0 = M^{-1}r_0$ (by linear CG), $p_0 = -q_0$;
2: **for** $i = 0, 1, \ldots$ **do**
3: compute α_i by (9), $x_{i+1} = x_i + \alpha_i p_i$, $r_{i+1} = r_i + \alpha_i AMp_i$;
4: if $\|r_{i+1}\|_1 / \|c\|_1 \leq$ `rtol`, **BREAK**;
5: $q_{i+1} = M^{-1}r_{i+1}$ by (linear) CG;
6: compute β_i by (9), $p_{i+1} = -q_{i+1} + \beta_i p_i$;
7: **end for**
8: **return** last x_i as an approximate solution.

Discussion

Both CG require solving $Mq = p$ in the inner iteration.

Another alternative is to rewrite symbolically

$$A = K^{-1/2} \underbrace{\left(K^{1/2} M K^{1/2} - \mu^2 I\right)}_{=:\widehat{A}} K^{1/2} = K^{-1/2} \widehat{A} K^{1/2}$$

$Ax = c$ is equivalent to $K^{-1/2} \widehat{A} K^{1/2} x = c \quad \Leftrightarrow \quad \widehat{A} \underbrace{K^{1/2} x}_{=:\hat{x}} = \underbrace{K^{1/2} c}_{=:\hat{c}}$.

$\widehat{A}$ is SPD because $0 < \mu < \lambda_1$. Can apply linear CG to $\widehat{A}\hat{x} = \hat{c}$ symbolically first and then translate to $Ax = c$.

Detail is omitted.

Convergence, Deflation (lock)

$(\rho_{\ell;j}, z_{\ell;j})$ is considered acceptable if $\dfrac{\|Hz_{\ell;j} - \rho_{\ell;j} z_{\ell;j}\|_2}{\|Hz_{\ell;j}\|_2 + |\rho_{\ell;j}| \, \|z_{\ell;j}\|_2} \leq \texttt{rtol}$.

Usually λ_j are converged to in order, i.e., the smallest eigenvalues emerge first.

Lock all acceptable approximate eigenpairs in $k_{\text{cvgd}} \times k_{\text{cvgd}}$ diagonal matrix $\boldsymbol{D}$ for eigenvalues and $2n \times k_{\text{cvgd}}$ tall matrix $\boldsymbol{Z}$ for eigenvectors.

Every time a converged eigenpair is detected, delete the converged $\rho_{\ell;j}$ and $z_{\ell;j}$ from Ω_ℓ and Z_ℓ, respectively, and expand $\boldsymbol{D}$ and $\boldsymbol{Z}$ to lock up the pair, accordingly.

At the same time, either reduce n_b by 1 or append a new column to Z to maintain n_b unchanged. The latter can be done by computing more than n_b eigenpairs at Line 9.

Deflate to avoid recomputing converged eigenpairs: Write $\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Y} \\ \boldsymbol{X} \end{bmatrix}$ and suppose $\boldsymbol{X}^T \boldsymbol{Y} = I_{k_{\text{cvgd}}}$.

- Modify K and M in form, but not explicitly, to $K + \zeta \boldsymbol{Y}\boldsymbol{Y}^T$ and $M + \zeta \boldsymbol{X}\boldsymbol{X}^T$, where ζ should be selected such that $\zeta + \lambda_1 \geq \lambda_{k_{\text{cvgd}}+n_b+1}$. Here we pre-assume the k_{cvgd} converged eigenpairs are indeed those for (λ_j, z_j) for $1 \leq j \leq k_{\text{cvgd}}$. This is usually so, but with no guarantee, of course.

Hyperbolic Quadratic Eigenvalue Problem

- Basics
- Rayleight Quotients
- Min-Max Principles
- SD and CG type Method

Hyperbolic $Q(\lambda) = A\lambda^2 + B\lambda + C$: Basics

$0 \prec A = A^{\mathsf{H}} \in \mathbb{C}^{n\times n}$, and $B = B^{\mathsf{H}}$, $C = C^{\mathsf{H}} \in \mathbb{C}^{n\times n}$.

$Q(\lambda)$ is *hyperbolic* if

$$(x^{\mathsf{H}}Bx)^2 - 4(x^{\mathsf{H}}Ax)(x^{\mathsf{H}}Cx) > 0 \quad \text{for } 0 \neq x \in \mathbb{C}^n.$$

This type Q arises, e.g., from dynamical systems that are overly damped.

Quadratic Eigenvalue Problem (QEP):

$$\text{find } \lambda \in \mathbb{C},\ 0 \neq x \in \mathbb{C}^n \text{ such that } Q(\lambda)x = 0.$$

λ: *quadratic eigenvalue*; x: *quadratic eigenvector*.

All quadratic eigenvalues of hyperbolic $Q(\lambda)$ are real:

$$\lambda_n^- \leq \cdots \leq \lambda_1^- < \lambda_1^+ \leq \cdots \leq \lambda_n^+.$$

Basics (cont'd)

For more basic properties of Hyperbolic QEP, see


C.-H. Guo and P. Lancaster. Algorithms for hyperbolic quadratic eigenvalue problems. *Math. Comp.*, 74:1777–1791, 2005.


Nicholas J. Higham, Françoise Tisseur, and Paul M. Van Dooren. Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems. *Linear Algebra Appl.*, 351-352:455–474, 2002.


A.S. Markus. *Introduction to the spectral theory of polynomial operator pencils*. Translations of mathematical monographs, vol. 71. AMS, Providence, RI, 1988.

Rayleigh Quotients

Given $x \neq 0$, consider

$$f(\lambda, x) := x^{\mathsf{H}} Q(\lambda) x = \lambda^2 (x^{\mathsf{H}} A x) + \lambda (x^{\mathsf{H}} B x) + (x^{\mathsf{H}} C x) = 0.$$

Always has two distinct real roots (as functions of x)

$$\rho_{\pm}(x) = \frac{-(x^{\mathsf{H}} B x) \pm \left[(x^{\mathsf{H}} B x)^2 - 4 (x^{\mathsf{H}} A x)(x^{\mathsf{H}} C x) \right]^{1/2}}{2 (x^{\mathsf{H}} A x)}.$$

Can show

$$\rho_+(x) \in [\lambda_1^+, \lambda_n^+], \quad \rho_-(x) \in [\lambda_n^-, \lambda_1^-].$$

Reasonable to define $\rho_{\pm}(x)$ as the *Rayleigh quotients* for the problem.

Courant-Fischer Type

Courant-Fischer type min-max principle (Duffin, 1955)

$$\lambda_i^+ = \max_{\operatorname{codim} \mathcal{X}=i-1} \min_{x\in\mathcal{X}} \rho_+(x), \quad \lambda_i^+ = \min_{\dim \mathcal{X}=i} \max_{x\in\mathcal{X}} \rho_+(x),$$
$$\lambda_i^- = \min_{\operatorname{codim} \mathcal{X}=i-1} \max_{x\in\mathcal{X}} \rho_-(x), \quad \lambda_i^- = \max_{\dim \mathcal{X}=i} \min_{x\in\mathcal{X}} \rho_-(x).$$

In particular,

$$\lambda_1^+ = \min_x \rho_+(x), \quad \lambda_n^+ = \max_x \rho_+(x),$$
$$\lambda_n^- = \min_x \rho_-(x), \quad \lambda_1^- = \max_x \rho_-(x).$$

- Duffin (1955) (though stated for hyperbolic Q with $B \succ 0$, $C \succ 0$, Duffin's proof works for the general hyperbolic case.)
- Also Markus (1988) (mostly about hyperbolic matrix polynomial of any degree), Voss (1982) (about certain nonlinear Q).

Trace Min/Max Type (Liang & Li, 2013)

$Q(\lambda) = A\lambda^2 + B\lambda + C$ hyperbolic. Its quadratic eigenvalues:

$$\lambda_n^- \leq \cdots \leq \lambda_1^- < \lambda_1^+ \leq \cdots \leq \lambda_n^+.$$

$k \leq n$, $X \in \mathbb{C}^{n\times k}$, $\operatorname{rank}(X) = k$. $X^{\mathrm{H}}Q(\lambda)X$ also hyperbolic.

Quadratic eigenvalues of $X^{\mathrm{H}}Q(\lambda)X$:

$$\lambda_{k,X}^- \leq \cdots \leq \lambda_{1,X}^- \leq \lambda_{1,X}^+ \leq \cdots \leq \lambda_{k,X}^+.$$

Trace Min/Max type principle

$$\inf_{\operatorname{rank}(X)=k} \sum_{j=1}^{k} \lambda_{j,X}^+ = \sum_{j=1}^{k} \lambda_j^+, \qquad \sup_{\operatorname{rank}(X)=k} \sum_{j=1}^{k} \lambda_{j,X}^+ = \sum_{j=1}^{k} \lambda_{n-k+j}^+,$$

$$\sup_{\operatorname{rank}(X)=k} \sum_{j=1}^{k} \lambda_{j,X}^- = \sum_{j=1}^{k} \lambda_j^-, \qquad \inf_{\operatorname{rank}(X)=k} \sum_{j=1}^{k} \lambda_{j,X}^- = \sum_{j=1}^{k} \lambda_{n-k+j}^-.$$

Corollary of a more general Wielandt type max/max principle (Liang & Li, 2013).

Cauchy Type

$Q(\lambda) = A\lambda^2 + B\lambda + C$ hyperbolic. Its quadratic eigenvalues:

$$\lambda_n^- \leq \cdots \leq \lambda_1^- < \lambda_1^+ \leq \cdots \leq \lambda_n^+.$$

$k \leq n$; $X \in \mathbb{C}^{k\times k}$, $\operatorname{rank}(X) = k$;

Quadratic eigenvalues of $X^{\mathrm{H}}Q(\lambda)X$:

$$\mu_k^- \leq \cdots \leq \mu_1^- < \mu_1^+ \leq \cdots \leq \mu_k^+.$$

Cauchy-type interlacing inequality

$$\begin{aligned} \lambda_i^+ \leq \mu_i^+ \leq \lambda_{i+n-k}^+, &\quad i = 1, \cdots, k, \\ \lambda_{j+n-k}^- \leq \mu_j^- \leq \lambda_j^-, &\quad j = 1, \cdots, k. \end{aligned}$$

- Veselić (2010).
- Also derivable from Wielandt type min-max principles (not presented here) (Liang & Li, 2013).

Rayleigh-Ritz Procedure for Hyperbolic Q

Recall two most important aspects in solving large scale eigenvalue problems: building good subspaces and seeking "*best possible*" approximations.

Given $\mathcal{Y} \in \mathbb{C}^n$ and $\dim \mathcal{Y} = m$, find the "*best possible*" approximations to some of $Q(\cdot)$'s quadratic eigenpairs "using $\mathcal{Y}$".

Can be done by a new "Rayleigh-Ritz" procedure. Let Y be $\mathcal{Y}$'s basis matrix.

Rayleigh-Ritz Procedure

1. Solve the QEP for $Y^{\mathsf{H}}Q(\lambda)Y$: $Y^{\mathsf{H}}Q(\mu_i^{\pm})Yy_i^{\pm} = 0$, where

$$\mu_m^- \le \cdots \le \mu_1^- < \mu_1^+ \le \cdots \le \mu_m^+.$$

2. Approximate quadratic eigenvalues: $\mu_i^{\pm} \approx \lambda_i^{\pm}$, approximate quadratic eigenvectors: $Yy_i^{\pm}$.

But in what sense and why are those $\mu_i^{\pm}$ and $Yy_i^{\pm}$ "*best possible*"?

Rayleigh-Ritz procedure for Hyperbolic Q (cont'd)

Trace Min/Max principle: $\inf_{\text{rank}(X)=k} \sum_{j=1}^{k} \lambda_{j,X}^{+} = \sum_{j=1}^{k} \lambda_j^{+}$ suggests that best possible approximations to λ_i^{+} $(1 \leq i \leq k)$ should be gotten so that

$$\sum_{j=1}^{k} \lambda_{j,X}^{+} \text{ is minimized, subject to } \text{span}(X) \subset \mathcal{Y},\ \text{rank}(X) = k.$$

The optimal value is $\sum_{j=1}^{k} \mu_j^{+}$.

Consequently, first few $\mu_i^{+} \approx \lambda_i^{+}$ are "*best possible*". Surprise: "interior" eigenvalues are usually hard to compute but this is not the case here.

Similarly to argue for last few $\mu_j^{+} \approx \lambda_{j+n-k}^{+}$ are "*best possible*".

Similarly to argue for first few $\mu_i^{-} \approx \lambda_i^{-}$ are "*best possible*". Surprise: "interior" eigenvalues are usually hard to compute but this is not the case here.

Similarly to argue for last few $\mu_j^{-} \approx \lambda_{j+n-k}^{-}$ are "*best possible*".

Gradients of Rayleigh quotients $\rho_\pm(x)$

Use $\rho(x)$ for either $\rho_+(x)$ or $\rho_-(x)$, and perturb x to $x+p$, $\|p\|$ tiny.
$\rho(x)$ is changed to $\rho(x+p) = \rho(x) + \eta + O(\|p\|^2)$. Then

$$[\rho(x)+\eta]^2 (x+p)^{\mathsf{H}} A(x+p) + [\rho(x)+\eta] (x+p)^{\mathsf{H}} B(x+p) + (x+p)^{\mathsf{H}} C(x+p) = 0$$

which gives, upon noticing $x^{\mathsf{H}} Q(\rho(x)) x = 0$, that

$$\begin{aligned}[2\rho(x)\, x^{\mathsf{H}} Ax + x^{\mathsf{H}} Bx]\eta + p^{\mathsf{H}}[\rho(x)^2 Ax + \rho(x) Bx + Cx] \\ + [\rho(x)^2 Ax + \rho(x) Bx + Cx]^{\mathsf{H}} p + O(\|p\|^2) = 0\end{aligned}$$

and thus

$$\eta = -\frac{p^{\mathsf{H}}[\rho(x)^2 Ax + \rho(x) Bx + Cx] + [\rho(x)^2 Ax + \rho(x) Bx + Cx]^{\mathsf{H}} p}{2\rho(x)\, x^{\mathsf{H}} Ax + x^{\mathsf{H}} Bx}.$$

Therefore the gradient of $\rho(x)$ at x is

$$\nabla\rho(x) = -\frac{2[\rho(x)^2 A + \rho(x) B + C]x}{2\rho(x)\, x^{\mathsf{H}} Ax + x^{\mathsf{H}} Bx}.$$

Important to notice that $\nabla\rho(x)$ is parallel to the residual vector

$$r_\pm(x) := [\rho_\pm(x)^2 A + \rho_\pm(x) B + C]x = Q(\rho_\pm(x))x.$$

Line Search

Steepest descent/ascent method for computing one of $\lambda_1^{\pm}$ can be readily given.

Fix two parameters "typ" and ℓ with varying ranges as

$$\text{typ} \in \{+, -\}, \quad \ell \in \{1, n\}$$

to mean that we are to compute the quadratic eigenpair $(\lambda_\ell^{\text{typ}}, u_\ell^{\text{typ}})$.

A key step of the method is the following line-search problem

$$t_{\text{opt}} = \underset{t \in \mathbb{C}}{\text{argopt}}\ \rho_{\text{typ}}(x + t\,p), \quad \text{argopt} = \begin{cases} \arg\min, & \text{for } (\text{typ}, \ell) \in \{(-, n), (+, 1)\}, \\ \arg\max, & \text{for } (\text{typ}, \ell) \in \{(-, 1), (+, n)\}. \end{cases}$$

where x is the current approximation to u_ℓ^{typ}, p is the search direction.

Not easy to do: Rayleigh quotient ρ_{typ} too complicated, unlike for (linear) eigenvalue problems.

Better way to solve by using min-max principle.

Line Search (cont'd)

Line Search is equivalent to find the best possible approximation within the subspace $\text{span}([x, p])$.

Suppose x and p are linearly independent and let $Y = [x, p]$.

Solve the 2-by-2 hyperbolic QEP for $Y^{\mathsf{H}} Q(\lambda) Y$ to get its quadratic eigenvalues

$$\mu_2^- \leq \mu_1^- < \mu_1^+ \leq \mu_2^+$$

and corresponding quadratic eigenvector $y_j^{\pm}$.

Table for selecting the next approximate quadratic eigenpair:

(typ, ℓ)	current approx.	next approx.
$(+, 1)$	$(\rho_+(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_1^+, Y y_1^+)$
$(+, n)$	$(\rho_+(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_2^+, Y y_2^+)$
$(-, 1)$	$(\rho_-(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_1^-, Y y_1^-)$
$(-, n)$	$(\rho_-(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_2^-, Y y_2^-)$

Steepest Descent/Ascent method

Basically it is Line Search along gradient direction.

Steepest Descent/Ascent method

Given an initial approximation $\boldsymbol{x}_0$ to u_ℓ^{typ}, and a relative tolerance `rtol`, the algorithm attempts to compute an approximate pair to $(\lambda_\ell^{\text{typ}}, u_\ell^{\text{typ}})$ with the prescribed `rtol`.

$\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|$, $\boldsymbol{\rho}_0 = \rho_{\text{typ}}(\boldsymbol{x}_0)$, $\boldsymbol{r}_0 = r_{\text{typ}}(\boldsymbol{x}_0)$;
for $i = 0, 1, \ldots$ **do**
  if $\|\boldsymbol{r}_i\|/(|\boldsymbol{\rho}_i|^2\|A\boldsymbol{x}_i\| + |\boldsymbol{\rho}_i|\,\|B\boldsymbol{x}_i\| + \|C\boldsymbol{x}_i\|) \le$ `rtol` **then**
    BREAK;
  else
    solve QEP for $Y_i^{\mathsf{H}} Q(\lambda) Y_i$, where $Y_i = [\boldsymbol{x}_i, \boldsymbol{r}_i]$;
    select the next approximate quadratic eigenpair $(\mu, y) = (\mu_j^{\text{typ}}, Y_i y_j^{\text{typ}})$ according to the table;
    $\boldsymbol{x}_{i+1} = y/\|y\|$, $\boldsymbol{\rho}_{i+1} = \mu$, $\boldsymbol{r}_{i+1} = r_{\text{typ}}(\boldsymbol{x}_{i+1})$;
  end if
end for
return $(\boldsymbol{\rho}_i, \boldsymbol{x}_i)$ as an approximate eigenpair to $(\lambda_\ell^{\text{typ}}, u_\ell^{\text{typ}})$.

Extended Steepest Descent/Ascent method

In Steepest Descent/Ascent method, the search space is spanned by

$$\boldsymbol{x}_i,\ \boldsymbol{r}_i = Q(\boldsymbol{\rho}_i)\boldsymbol{x}_i.$$

It is the second order Krylov subspace $\mathcal{K}_2(Q(\boldsymbol{\rho}_i), \boldsymbol{x}_i)$ of $Q(\boldsymbol{\rho}_i)$ on $\boldsymbol{x}_i$.

One way to improve the method is to use a higher order Krylov subspace

$$\mathcal{K}_m(Q(\boldsymbol{\rho}_i), \boldsymbol{x}_i) = \mathrm{span}\{\boldsymbol{x}_i, Q(\boldsymbol{\rho}_i)\boldsymbol{x}_i, \ldots, [Q(\boldsymbol{\rho}_i)]^{m-1}\boldsymbol{x}_i\}.$$

Let Y_i be a basis matrix of $\mathcal{K}_m(Q(\boldsymbol{\rho}_i), \boldsymbol{x}_i)$. Solve m-by-m hyperbolic QEP for $Y_i^{\mathrm{H}} Q(\lambda) Y_i$ to get its quadratic eigenvalues

$$\mu_m^- \le \cdots \le \mu_1^- < \mu_1^+ \le \cdots \le \mu_m^+$$

and corresponding quadratic eigenvectors $y_j^{\pm}$.

Table for selecting the next approximate quadratic eigenpair:

(typ, ℓ)	current approx.	next approx.
$(+, 1)$	$(\rho_+(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_1^+, Y_i y_1^+)$
$(+, n)$	$(\rho_+(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_m^+, Y_i y_m^+)$
$(-, 1)$	$(\rho_-(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_1^-, Y_i y_1^-)$
$(-, n)$	$(\rho_-(\boldsymbol{x}), \boldsymbol{x})$	$(\mu_m^-, Y_i y_m^-)$

Extended Steepest Descent/Ascent method (cont'd)

Extended Steepest Descent/Ascent method

Given an initial approximation $\boldsymbol{x}_0$ to u_ℓ^{typ}, and a relative tolerance `rtol`, and the search space dimension m, the algorithm attempts to compute an approximate pair to $(\lambda_\ell^{\mathrm{typ}}, u_\ell^{\mathrm{typ}})$ with the prescribed `rtol`.

$\boldsymbol{x}_0 = \boldsymbol{x}_0/\|\boldsymbol{x}_0\|$, $\boldsymbol{\rho}_0 = \rho_{\mathrm{typ}}(\boldsymbol{x}_0)$, $\boldsymbol{r}_0 = r_{\mathrm{typ}}(\boldsymbol{x}_0)$;
for $i = 0, 1, \ldots$ **do**
  if $\|\boldsymbol{r}_i\|/(|\boldsymbol{\rho}_i|^2\|A\boldsymbol{x}_i\| + |\boldsymbol{\rho}_i|\,\|B\boldsymbol{x}_i\| + \|C\boldsymbol{x}_i\|) \le$ `rtol` **then**
    BREAK;
  else
    compute a basis matrix Y_i for $\mathcal{K}_m(Q(\boldsymbol{\rho}_i), \boldsymbol{x}_i)$;
    solve QEP for $Y_i^{\mathsf{H}} Q(\lambda) Y_i$ to get its quadratic eigenvalues $\mu_j^{\pm}$ and eigenvectors $y_j^{\pm}$;
    select the next approximate quadratic eigenpair $(\mu, y) = (\mu_j^{\mathrm{typ}}, Y y_j^{\mathrm{typ}})$ according to the table;
    $\boldsymbol{x}_{i+1} = y/\|y\|$, $\boldsymbol{\rho}_{i+1} = \mu$, $\boldsymbol{r}_{i+1} = r_{\mathrm{typ}}(\boldsymbol{x}_{i+1})$;
  end if
end for
return $(\boldsymbol{\rho}_i, \boldsymbol{x}_i)$ as an approximate eigenpair to $(\lambda_\ell^{\mathrm{typ}}, u_\ell^{\mathrm{typ}})$.

Rate of Convergence

Rate of Convergence (Liang & Li, 2013)

$$|\boldsymbol{\rho}_{i+1} - \lambda_\ell^{\mathrm{typ}}| \le \varepsilon_m^2 |\boldsymbol{\rho}_i - \lambda_\ell^{\mathrm{typ}}| + O\big(\varepsilon_m |\boldsymbol{\rho}_i - \lambda_\ell^{\mathrm{typ}}|^{3/2} + |\boldsymbol{\rho}_i - \lambda_\ell^{\mathrm{typ}}|^2\big),$$

where

$$\varepsilon_m = \min_{g \in \mathbb{P}_{m-1}, g(\sigma_1) \neq 0} \max_{i \neq 1} \frac{|g(\sigma_i)|}{|g(\sigma_1)|},$$

and σ_j for $1 \le j \le n$ are eigenvalues of $Q(\boldsymbol{\rho}_i)$ arranged as in

$$\begin{aligned} &\sigma_1 > 0 > \sigma_2 \ge \cdots \ge \sigma_n \quad \text{if } (\mathrm{typ}, \ell) \in \{(+,1),(-,1)\}, \quad \text{or,} \\ &\sigma_1 < 0 < \sigma_2 \le \cdots \le \sigma_n \quad \text{if } (\mathrm{typ}, \ell) \in \{(+,n),(-,n)\}. \end{aligned}$$

- While the result is similar to the one for $A - \lambda B$ ($B \succ 0$), it is *much much* more complicated to prove.
- Important: rate depends on eigenvalue distribution of $Q(\boldsymbol{\rho}_i)$. Shed light to preconditioning:

  $$Q(\boldsymbol{\rho}_i) \approx L_i D_i L_i^{\mathsf{H}}, \quad D_i = \mathrm{diag}(\pm 1),$$

  and use Extended Steepest Descent/Ascent method on $L^{-1} Q(\lambda) L^{-\mathsf{H}}$.

  Should reformulate for implementation sake. Detail omitted.

CG methods, Block Variations

Straightforward applications of ideas presented for $A - \lambda B$ earlier.
Left as exercises ...