Numerical solution of matrix eigenvalue problems
Part 3: Preconditioned Eigensolvers

Daniel Kressner
ETH Zürich

kressner@math.ethz.ch
http://www.math.ethz.ch/~kressner

ZSS 2008

Outline

- Preconditioned inverse iteration
- CG
- LOPCG
- Jacobi-Davidson

Setting: A spd but does not admit exact inverse. Eigenvalues

$$\lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n, \quad \frac{\lambda_2 - \lambda_1}{\lambda_n - \lambda_2} \approx 0.$$

Want to compute λ_1. Convergence of Lanczos slow.

Basic idea: replace A^{-1} by B^{-1} with preconditioner $B \approx A$.

But: If immediately plugged into power method $\rightsquigarrow$ smallest eigenvalue of B instead of A.

Inject information on A via residuum.

$$r_i = Av_i - \mu_i v_i, \quad \mu_i = v_i^H A v_i, \quad \|v_i\|_2 = 1.$$

Inexact solution of refinement equation $A\triangle v_i = -r_i \rightsquigarrow$

$$\begin{aligned} \triangle v_i &= -B^{-1}(Av_i - \mu_i v_i), \\ \widetilde{v}_{i+1} &= v_i + \triangle v_i = v_i - B^{-1}(Av_i - \mu_i v_i) \\ v_{i+1} &= \widetilde{v}_{i+1} / \|\widetilde{v}_{i+1}\|_2. \end{aligned}$$

Convergence analysis

Interpretation as perturbed inverse iteration:

$$\widetilde{v}_{i+1} = \mu_i A^{-1} v_i + (I - B^{-1}A)(v_i - \mu_i A^{-1} v_i)$$

Show ppd.m.

Theorem [Neymeyr]: If $\|I - B^{-1}A\|_A \leq \gamma < 1$ then the sequence μ_i decreases monotonically and (μ_i, v_i) converges to an eigenpair of A (usually to the smallest). Provided $\mu_i \to \lambda_1$, we have

$$\frac{\mu_{i+1} - \lambda_1}{\mu_i - \lambda_1} \leq \frac{1 - (1-\gamma^2)\frac{\lambda_2 - \mu_i}{\lambda_2}}{1 + (1-\gamma)^2 \frac{(\mu_i - \lambda_1)(\lambda_2 - \mu_i)}{\lambda_1 \lambda_2}} < 1.$$

$\rightsquigarrow$ Convergence rate asymptotically bounded by

$$\frac{(1-\gamma^2)\lambda_1 + \gamma^2 \lambda_2}{\lambda_2}.$$

$\rightsquigarrow$ Convergence rate independent of λ_n!

Consider symmetric positive definit A. Then

$$\lambda_1 = \min_{x\neq 0} \rho(x), \quad \rho(x) := \frac{x^T A x}{x^T x}.$$

Basic idea: Given previouse iterate v_k and search direction p_k choose new iterate

$$x_{k+1} = x_k + \delta_k p_k$$

such that $\rho(x_{k+1}) \to \min$. Strictly convex problem $\rightsquigarrow$ step size δ_k is uniquely determined.

Choice of search direction as in CG for linear systems:

$$p_k \perp_A p_{k-1}.$$

Convergence rate: $(\sqrt{\kappa}-1)/(\sqrt{\kappa}+1)$ with

$$\kappa = (\lambda_n - \lambda_1)/(\lambda_2 - \lambda_1).$$

Depends on λ_n!

LOPCG

Wanttoincorporatepreconditioner$B^{-1} \approx A^{-1}$. Idea: Throw in all available (local) information to minimize Rayleigh quotient:

$$\min_{\alpha\in\mathbb{R},\beta\in\mathbb{R}} \rho(B^{-1}r_i + \alpha v_{i-1} + \beta v_i) \to \min.$$

- α, β can be determined by choosing minimal Ritz vector of A w.r.t. $\text{span}\{B^{-1}r_i, v_{i-1}, v_i\}$;
- not worse than preconditioned inverse iteration; show runlopcg
- numerical experiments reported by [Knyazev'01] reveal little benefit when taking "older" v_i into account;
- straightforward to generalize to subspaces $\rightsquigarrow$ LOBPCG.

Basic idea: Temporarily assume exact knowledge of (simple) eigenvalue λ of A. Correction equation for approximate eigenvector u with $\|u\|_2 = 1$:

$$A(u+v) = \lambda(u+v), \quad u \perp v. \tag{1}$$

Constraint $u \perp v \rightsquigarrow$ unique solvability; $u+v$ is exact eigenvector!

Components in the direction of u are irrelevant in (1) $\rightsquigarrow$

$$(I - uu^H)(A - \lambda I)(I - uu^H)v = -(A - \mu I)u =: -r, \quad \mu = u^H A u.$$

Replace unknown λ by $\mu \rightsquigarrow$ Jacobi-Davidson correction equation

$$(I - uu^H)(A - \mu I)(I - uu^H)v = -r.$$

Closely related to Rayleigh quotient iteration if correction equation is solved exactly:

$$u + v = \alpha(A - \mu I)^{-1}u.$$

$\rightsquigarrow$ local quadratic/cubic convergence.

JD: Subspace acceleration

Input: start vector u_1 with $\|u_1\|_2 = 1$, matrix A, integer k.
Output: Approximate eigenpair (μ_k, u_k).

Set $\mu_1 = u_1^H A u_1$.
for $j = 1, 2, \ldots, k-1$ **do**
  $r = Au_j - \mu_j u_j$
  Solve correction equation $(I - u_j u_j^H)(A - \mu_j I)(I - u_j u_j^H)v_j = -r$.
  $u_{j+1} = (u_j + v_j)/\|u_j + v_j\|_2$
  $\mu_{j+1} = u_{j+1}^H A u_{j+1}$
end for

Similar idea as power method $\rightarrow$ Krylov subspace: Collect all previously computed corrections in subspace:

$$\mathcal{V}_j = \operatorname{span}\{u_1, v_2, v_3, \ldots, v_j\}.$$

and choose corresponding Ritz pair in next iteration.

Input: start vector u_1 with $\|u_1\|_2 = 1$, matrix A, integer k, target τ.
Output: Approximate eigenpair (μ_k, u_k).

Set $\mu_1 = u_1^H A u_1$, $V_1 = [u_1]$.
for $j = 1, 2, \dots, k-1$ **do**
  $r = Au_j - \mu_j u_j$
  Solve correction equation $(I - u_j u_j^H)(A - \mu_j I)(I - u_j u_j^H) v_{j+1} = -r$.
  $\widetilde{v}_{j+1} = (I - V_j V_j)^H v_{j+1}$, $v_{j+1} \leftarrow \widetilde{v}_{j+1} / \|\widetilde{v}_{j+1}\|_2$, $V_{j+1} = [V_j, v_{j+1}]$.
  Compute $j+1$ Ritz values as eigenvalues of $H_{j+1} = V_{j+1}^H A V_{j+1}$.
  Set μ_{j+1} to Ritz value closest to τ.
  Set u_{j+1} to corresponding Ritz vector.
end for

Remarks:

- H_j is *not* a Hessenberg matrix. H_{j+1} can be obtained cheaply from H_j (by bordering) provided that AV_j is maintained.
- Reorthogonalization of v_{j+1} may be necessary.
- To avoid erratic global convergence, set $\mu_j = \tau$ during the first few iterations.
- Alternative to Ritz vectors: Refined and harmonic Ritz vectors.

JD: Inexact solves

Major feature of JD: robust convergence even with (very) inexact solves of correction equation. Apply GMRES or MINRES to

$$(I - uu^H)(A - \mu I)(I - uu^H) v = -r, \quad v \perp u. \tag{2}$$

Since $r \perp u$, Krylov subspace method will not "see" singularity of

$$\widetilde{A} := (I - uu^H)(A - \mu I)(I - uu^H)$$

Applying a preconditioner B directly to (2) would destroy this property. Need to (formally) use modified preconditioner

$$\widetilde{B} = (I - uu^H) B (I - uu^H).$$

- $\widetilde{B} x = y$ has unique solution $x \perp u$ provided $y \perp u$.
- Denote the solution operator by $\widetilde{B}^\dagger : y \mapsto x$ (pseudoinverse) $\rightsquigarrow$ preconditioned correction equation

$$\widetilde{B}^\dagger \widetilde{A} v = -\widetilde{B}^\dagger r.$$

Typically only B^{-1} and *not* $\widetilde{B}^\dagger$ is available. To compute $\widetilde{B}^\dagger y$ for some $y \perp u$:

$$\widetilde{B}^\dagger y = B^{-1}y - \alpha B^{-1}u, \quad \alpha = \frac{u^H B^{-1} y}{u^H B^{-1} u}.$$

Further remarks:

- Restarts are simple: compress V_j to Ritz vectors of interest.
- If several eigenvalues are needed: Collect converged f Schur vectors in basis Q_f, replace correction equation by

$$(I - uu^H)(I - Q_f Q_f^H)(A - \mu I)(I - Q_f Q_f^H)(I - uu^H)v = -r,$$

  with the constraints $v \perp u$, $v \perp \mathrm{span}(Q_f)$. Extend search basis V_k to $[Q_f, V_k]$. $\rightsquigarrow$ JDQR.

Run jdqr.m

Literature

- The definitive guide to the preconditioned inverse iteration is Neymeyr's series of papers on this subject.
- The description of CG and LOPCG is taken from [Templates for the Solution of Algebraic Eigenvalue Problems. SIAM, 2000]. A more detailed description can be found in Knyazev's original paper.
- The description of Jacobi-Davidson is taken from [Peter Arbenz. Lecture notes on solving large-scale eigenvalue problems].
- A derivation of Jacobi-Davidson from the Newton method can be found in [G. W. Stewart. Matrix Algorithms. Vol. II. 2001].