Iterative methods for symmetric eigenvalue problems

Imre Pólik, PhD

McMaster University
School of Computational Engineering and Science

February 11, 2008

Outline

1. The power method and its variants
   - Power method
   - Inverse power method
   - Rayleigh quotient iteration
   - Subspace iteration method
2. Krylov subspace methods
   - Rayleigh-Ritz method
   - Lanczos method
   - Implicitly restarted Lanczos method
   - Band Lanczos method
3. Jacobi-Davidson method

Power method

- Find the dominating eigenvalue/eigenvector

$$
\begin{aligned}
v_{k+1} &= y_k / \left\| y_k \right\|_2 \\
y_{k+1} &= A v_{k+1} \\
\lambda_{k+1} &= v_{k+1}^T y_{k+1}
\end{aligned}
$$

- Only multiplication is involved
- Converges unless $v_0 \perp v_{\max}$
- Convergence rate: $|\lambda_{\max} / \lambda_{\max-1}|$
- Problems
  - multiple/close largest eigenvalues
  - only the largest eigenvalue is computed

Inverse power method

- Inner eigenvalues: $\lambda\left((A-\sigma I)^{-1}\right) = \frac{1}{\lambda(A)-\sigma}$
- Apply the power method to $(A-\sigma I)^{-1}$

$$
\begin{aligned}
v_{k+1} &= y_k / \|y_k\|_2 \\
y_{k+1} &= (A-\sigma I)^{-1} v_{k+1} \\
\lambda_{k+1} &= v_{k+1}^T y_{k+1}
\end{aligned}
$$

- Converges to the dominating eigenvalue of $(A-\sigma I)^{-1}$
- Converges unless $v^0 \perp v_{\max}$
- Convergence rate: $|(\lambda_{\max}-\sigma)/(\lambda_{\max-1}-\sigma)|$, linear
- Viable only if $(A-\sigma I)y = v$ is easily solvable

Rayleigh quotient iteration

- Change the shift in each iteration

$$v_{k+1} = y_k / \|y_k\|_2$$
$$\sigma_{k+1} = v_{k+1}^T A v_{k+1} / \|v_{k+1}\|^2$$
$$y_{k+1} = (A - \sigma_{k+1} I)^{-1} v_{k+1}$$
$$\lambda_{k+1} = v_{k+1}^T y_{k+1}$$

- Convergence properties are unclear
  - Finds *an* eigenvalue faster than inverse iteration
  - Cubic convergence
  - Does not necessarily find $\lambda_{\max}$
  - May not converge to an eigenvalue
- $A - \sigma_{k+1} I$ will become singular
- New factorization in every iteration

Subspace iteration method

- Invariant subspaces are robust, eigenvectors are not

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & \varepsilon \\ 0 & \varepsilon & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1+\varepsilon \end{pmatrix}$$

- It is better to identify the invariant subspaces

$$\begin{aligned} \text{QR factorize} \quad Y_k &= V_{k+1}R_{k+1} \\ Y_{k+1} &= AV_{k+1} \\ H_{k+1} &= V_{k+1}^T Y_{k+1} \end{aligned}$$

- $Y, V \in \mathbb{R}^{n\times p}$, $H \in \mathbb{R}^{p\times p}$
- The eigenvalues of H are the largest eigenvalues of A
- Clustered (not multiple) eigenvalues
- Choosing p smartly
- Can also be applied to $(A - \sigma I)^{-1}$
- Software: EA12 in HSL

Krylov subspace methods

- Problems with power iteration based methods
  - Extremal eigenvalues only
  - Internal eigenvalues require solution of a linear system
  - $A^k v$ is used as the best guess for an eigenvector
- Krylov subspace: $\operatorname{span}\left\{v, Av, \ldots, A^k v\right\}$
- Find the best approximate eigenvector (Ritz vectors)
- Columns of Q_k are orthogonal, span Krylov space
- $\lambda(Q_k^T A Q_k)$ approximates $\lambda(A)$
- Choose a Q_k to simplify the structure of $\underbrace{Q_k^T A Q_k}_{T_k}$

Lanczos method

- Gradually build the Krylov subspace
- Maintain an orthogonal basis Q_k, T_k tridiagonal
- Find the corresponding Ritz vectors

$v_0 = 0, \beta_1 = 0, v_1$ random unit

repeat

$$q_j = Av_j - \beta_j v_{j-1}$$
$$\alpha_j = q_j^T v_j$$
$$q_j = q_j - \alpha_j v_j$$
$$\beta_{j+1} = \|q_j\|$$
$$v_{j+1} = q_j / \beta_{j+1}$$

- Extreme eigenvalues converge first
- Can also be applied to $(A - \sigma I)^{-1}$
- Memory consumption increases

Implicitly restarted Lanczos method

- Prevents k growing too much
- Applies shifts μ_i to the algorithm
- Equivalently, changes v_0
- How to choose the shifts?
- Flexible eigenvalue configurations
- Locking/purging eigenvalues
- Software: `ARPACK` (also in Matlab)

Band Lanczos method

- Multiple starting vectors, finds more eigenvalues $\operatorname{span}\left\{V, AV, \ldots, A^k V\right\}$
- Suitable for multiple/clustered eigenvalues
- T is block tridiagonal

Jacobi-Davidson method

- Problems with Lanczos
  - only efficient if the eigenvalues are well separated
  - needs $(A - \sigma I)^{-1} y$ for internal eigenvalues
- Build a different set of orthogonal vectors spanning the Galerkin vectors
- Interior eigenvalues without inversion
- Very good if A has multiple eigenvalues
- Software: JDQR (Matlab)

Z. 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.

James W. Demmel.
Applied Numerical Linear Algebra.
SIAM, Philadelphia, 1997.