Numerical solution of matrix eigenvalue problems
Part 1: Power method and friends

Daniel Kressner
ETH Zürich

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

ZSS 2008

Outline

- power method
- subspace iteration
- inverse iteration
- Rayleigh-quotient iteration

The power method

Input: start vector $v_0 \in \{\mathbb{R}, \mathbb{C}\}^n$, matrix $A \in \{\mathbb{R}, \mathbb{C}\}^{n \times n}$.

for $i = 0, 1, 2, \ldots$ **do**

$v_{i+1} = Av_i$

end for

Let $\lambda_1, \ldots, \lambda_n$ be eigenvalues of A s.t.

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$$

with corresponding eigenvectors $x_1, x_2, \ldots, x_n$.

Power method converges to x_1:

$$\theta(x_1, v_i) = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^i\right)$$

Angle between vectors $v, w \in \mathbb{C}^n$:

$$\cos\theta(v, w) = \frac{|w^H v|}{\|v\|_2 \|w\|_2}$$

The power method

Input: start vector $v_0 \in \{\mathbb{R}, \mathbb{C}\}^n$, matrix $A \in \{\mathbb{R}, \mathbb{C}\}^{n \times n}$.
Output: Approximation v_i to dominant eigenvector.

for $i = 0, 1, 2, \ldots$ **do**
  $\widetilde{v}_{i+1} = Av_i$
  $v_{i+1} = \widetilde{v}_{i+1} / \|\widetilde{v}_{i+1}\|_2$
end for

To get an approximation to λ_1, choose μ_i s.t.

$$\|Av_i - \mu_i v_i\|_2 \to \min$$

$\rightsquigarrow$ Rayleigh quotient $\mu_i := v_i^H A v_i = v_i^H \widetilde{v}_{i+1}$.

$$|\mu_i - \lambda_1| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^i \right).$$

For symmetric A:

$$|\mu_i - \lambda_1| \leq |\lambda_n - \lambda_1| \sin^2 \theta(x_1, v_0) \left| \frac{\lambda_2}{\lambda_1} \right|^{2k}.$$

The power method

Termination? Is $|\mu_{i+1} - \mu_i| \approx 0$ a suitable criterion?

- For symmetric/normal matrices, yes:

$$|\mu_{i+1} - \mu_i| \le \text{tol} \quad \Rightarrow \quad |\mu_k - \lambda| \le \frac{1}{1 - |\lambda_2/\lambda_1|}\text{tol}.$$

- For nonnormal matrices, no. Transient growth! Show transient.m

  Only reasonable choice: $r := Av_i - \mu_i v_i \approx 0$.

  Backward error interpretation: (μ_i, v_i) *exact* eigenvalue/eigenvector pair of *perturbed* matrix

$$A + \triangle A, \quad \triangle A := -r v_i^H.$$

  With this choice, $\|\triangle A\|_2 = \|r\|_2$.

Snippets of perturbation analysis

First-order influence of $\triangle A$ on accuracy of eigenvalues/eigenvectors:

$$Ax = \lambda x, \quad \|x\|_2 = 1.$$

$$\Downarrow$$

$$(A + \epsilon \triangle A) x(\epsilon) = \lambda(\epsilon) x(\epsilon), \quad \|x(\epsilon)\|_2 = 1.$$

$$\Downarrow$$

$$\dot{\lambda}(0) = \frac{y^H \triangle A x}{y^H x},$$

where $y^H A = \lambda y^H$.

Accuracy of eigenvalues/eigenvectors

In general:

$$|\mu_i - \lambda_1| \le \frac{1}{|y_1^H x_1|}\|\triangle A\|_2 + O(\|\triangle A\|_2^2)$$

and

$$\theta(v_i, \theta_1) \le \|[(I - x_1 x_1^H)(A - \lambda_1 I)(I - x_1 x_1^H)]^\dagger\|_2 \|\triangle A\|_2 + O(\|\triangle A\|_2^2).$$

If A is symmetric:

$$|\mu_i - \lambda_1| \le \|\triangle A\|_2.$$

and

$$\sin\theta(x_1, v_i) = \frac{1}{|\mu_i - \lambda_2|}\|\triangle A\|_2.$$

A posteriori estimates

- If A is nonsymmetric, apply power method to A and A^H. Obtain approximations v_i and w_i to right/left eigenvectors x_1, y_1. Then

$$|\mu_i - \lambda_1| \approx \frac{1}{|w_i^H v_i|} \max\{\|Av_i - \mu_i v_i\|_2, \|w_i^H A - \mu_i w_i^H\|_2\}.$$

- If A is symmetric,

$$|\mu_i - \lambda_i| \le \|r\|_2, \quad \sin\theta(x_1, v_i) = \frac{1}{|\mu_i - \lambda_2|}\|r\|_2.$$

- Alternative for symmetric A: Kato-Temple inequality. Assume $\exists$ interval (a, b) containing μ_i, λ_1 and no other eigenvalue of A. Then

$$-\frac{\|r\|_2^2}{\mu_i - a} \le \mu_i - \lambda_i \le \frac{\|r\|_2^2}{b - \mu_i}.$$

Subspace iteration

Natural generalization of power method from vectors to subspaces:

Input: Start matrix $V_0 \in \{\mathbb{R}, \mathbb{C}\}^{n\times k}$, matrix $A \in \{\mathbb{R}, \mathbb{C}\}^{n\times n}$.
for $i = 0, 1, 2, \ldots$ **do**
$\widetilde{V}_{i+1} = AV_i$
Compute QR decomposition $\widetilde{V}_{i+1} = V_{i+1}R_{i+1}$.
end for

Convergence: Consider two groups of eigenvalues

$$|\lambda_1| \geq \cdots \geq |\lambda_k| > |\lambda_{k+1}| \geq \cdots \geq |\lambda_n|$$

with corresponding invariant subspaces $\mathcal{X}_1$, $\mathcal{X}_2$.

If $\theta(\mathcal{X}_1, \mathcal{V}_0) < \pi/2$, then

$$\theta(\mathcal{X}_1, \mathcal{V}_i) = O\left(\left|\frac{\lambda_{k+1}}{\lambda_k}\right|^i\right)$$

Angles between subspaces:

$$\sin\theta(\mathcal{V}, \mathcal{W}) := \min_{\substack{v\in\mathcal{V}\\ \|v\|_2=1}} \min_{w\in\mathcal{W}} \|v-w\|_2.$$

Given ONBs V, W of $\mathcal{V}, \mathcal{W}$:

$$\cos\theta(\mathcal{V}, \mathcal{W}) = \|W^H V\|_2.$$

Subspace iteration

Eigenvalue/eigenvector extraction from Galerkin condition

$$Av - \mu v \perp \mathcal{V}_i, \quad v \in \mathcal{V}_i.$$

This implies

- μ is an eigenvalue of the $k \times k$ matrix $V_i^H A V_i$ (compression of A);
- $v = V_i w$ with w, $\|w\|_2 = 1$, eigenvector of $V_i^H A V_i$ belonging to μ.

μ, v are called Ritz value/Ritz vector.

In the ith iteration, there are k such Ritz pairs

$$(\mu_i^{(1)}, v_i^{(1)}), (\mu_i^{(2)}, v_i^{(2)}), \ldots, (\mu_i^{(k)}, v_i^{(k)}).$$

The space $\mathcal{V}_i = \text{span}\{v_i^{(1)}, \ldots, v_i^{(k)}\}$ converges as a whole with rate $|\lambda_{k+1}/\lambda_k|$. However, some $v_i^{(j)}$ may converge significantly faster than others. Show ssiter.m.

Let A be symmetric and order eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n.$$

Minimax theorems:

$$\begin{aligned} \lambda_j &= \max_{\dim(\mathcal{V})=j} \min_{\substack{v \in \mathcal{V} \\ \|v\|_2=1}} v^T A v, \\ \lambda_{n-j+1} &= \min_{\dim(\mathcal{V})=j} \max_{\substack{v \in \mathcal{V} \\ \|v\|_2=1}} v^T A v. \end{aligned}$$

As a consequence, Ritz values interlace eigenvalues of A:

$$\lambda_j \geq \mu^{(j)} \geq \lambda_{n-j+1},$$

$\rightsquigarrow$ Ritz values approach eigenvalues from below.

Inverse iteration

Eigenvalue closest to $\tau \in \mathbb{C}$ is dominant eigenvalue of

$$(A - \tau I)^{-1}.$$

Input: start vector v_0, matrix A, target $\tau \in \mathbb{C}$.
Output: Approximation (μ_i, v_i) to eigenpair closest to τ.

for $i = 0, 1, 2, \ldots$ **do**
  Solve $(A - \tau I)\widetilde{v}_{i+1} = v_i$
  $v_{i+1} = \widetilde{v}_{i+1} / \|\widetilde{v}_{i+1}\|_2$
  $\mu_{i+1} = (\widetilde{v}_{i+1}^H v_{i+1})^{-1} + \tau$
end for

Not suited for inexact application of $(A - \tau I)^{-1}$!

Inverse iteration

Order eigenvalues

$$|\lambda_1 - \tau| \geq \cdots \geq |\lambda_{n-1} - \tau| > |\lambda_n - \tau|.$$

Convergence to (λ_n, x_n):

$$\begin{aligned} |\mu_i - \lambda_n| &= O\left(\left|\frac{\lambda_n - \tau}{\lambda_{n-1} - \tau}\right|^i\right) \\ \theta(x_n, v_i) &= O\left(\left|\frac{\lambda_n - \tau}{\lambda_{n-1} - \tau}\right|^i\right) \end{aligned}$$

Rayleigh-quotient iteration

Tempting idea: Improve convergence by adjusting τ.

Input: start vector v_0 with $\|v_0\|_2 = 1$, matrix A.
Output: Approximation (μ_i, v_i) to some eigenpair.

for $i = 0, 1, 2, \ldots$ **do**
    Set $\tau = v_i^H A v_i$.
    Solve $(A - \tau I)\widetilde{v}_{i+1} = v_i$
    $v_{i+1} = \widetilde{v}_{i+1} / \|\widetilde{v}_{i+1}\|_2$
    $\mu_{i+1} = v_{i+1}^H A v_{i+1}$
end for

Local convergence quadratic or even cubic (when A normal).

Global convergence critical!

Start rqiglobal.m

Grassmann Rayleigh-quotient iteration

Natural generalization to subspaces.

Input: start ONB V_0, symmetric matrix A.
Output: Approximate ONB V_i to some invariant subspace.

for $i = 0, 1, 2, \ldots$ **do**
  Set $T = V_i^H A V_i$.
  Solve Sylvester equation $A\widetilde{V}_{i+1} - \widetilde{V}_{i+1}T = V_i$.
  Compute ONB V_{i+1} of $\widetilde{V}_{i+1}$.
end for

Local convergence cubic. Advantageous over RQI, when several eigenvalues are needed:

- Avoid repeated convergence to same eigenvalues.
- In the limit, effective conditioning of $X \mapsto AX - XT$ only depends on external gap.
- Domain of attraction primarily depends on external gap.

Literature

- Virtually every book on numerical linear algebra covers the power method. The subspace iteration is covered, e.g., in [G. H. Golub and C. F. Van Loan. Matrix Computations. 1996], [G. W. Stewart. Matrix Algorithms. Vol. II. 2001].
- An implementation of the subspace iteration can be found in [Z. Bai and G. W. Stewart. Algorithm 776. SRRIT – A FORTRAN subroutine to calculate the dominant invariant subspaces of a nonsymmetric matrix. ACM TOMS, 23:494–513, 1998].
- The standard reference on the perturbation of *nonsymmetric* eigenvalue problems is [G. W. Stewart and J.-G. Sun. Matrix Perturbation Theory. Academic Press, New York, 1990].
- An overview of available relative perturbation bounds for *symmetric* eigenvalue problems is given in [I. C. F. Ipsen. Relative Perturbation Bounds for Matrix Eigenvalues and Singular Values. Acta Numerica, pp 151–201, (1998)].
- The Grassmann RQI and related methods are described in the recent monograph [P.-A. Absil, R. Mahony, R. Sepulchre. Optimization Algorithms on Matrix Manifolds. Princeton University Press, 2008].