Recitation 6

Recitation Instructor: Shivam Verma
Email: shivamverma@nyu.edu
Ph: 718-362-7836
Office hours: WWH 605 (2.50 - 4.50 pm, Tuesdays)

Brief Overview

- Gershgorin's disc theorems/Sturm sequence (review)
- QR algorithm
- Inverse iteration
- Rayleigh coefficient method
- Perturbation analysis and condition numbers

Gershgorin's theorems (review)

Theorem 1:

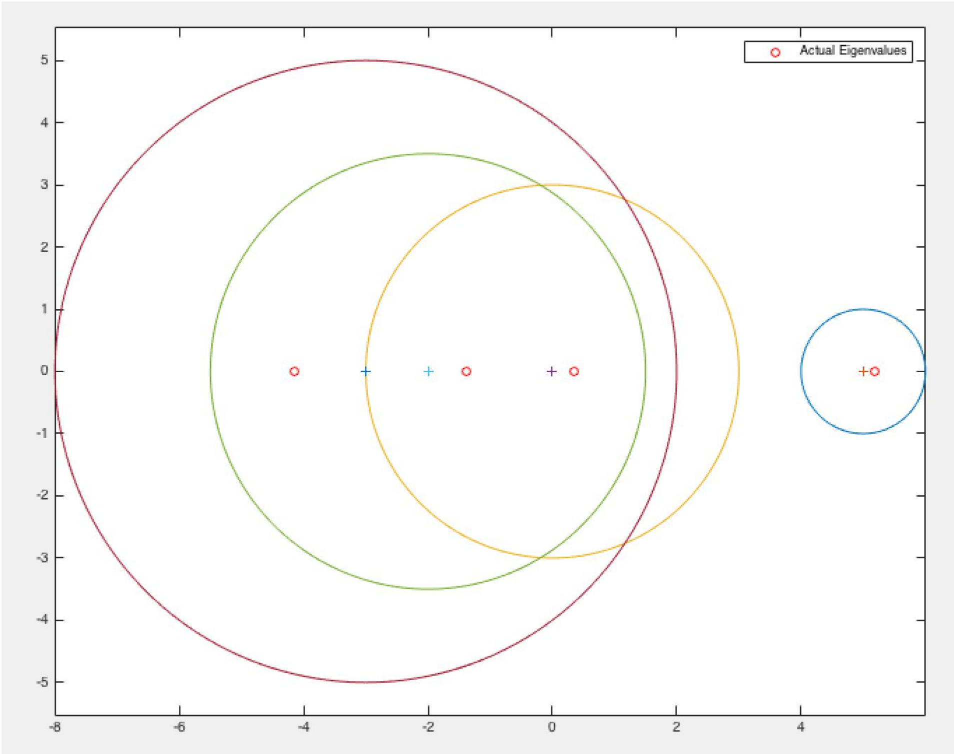
Every eigenvalue of matrix A_{nn} satisfies:

$$|\lambda - A_{ii}| \leq \sum_{j \neq i} |A_{ij}| \qquad i \in \{1, 2, ..., n\}$$

Every eigenvalue of a matrix A must lie in a Gershgorin disc corresponding to the columns of A.

Theorem 2:

A Subset G of the Gershgorin discs is called a disjoint group of discs if no disc in the group G intersects a disc which is not in G. If a disjoint group G contains r nonconcentric discs, then there are r eigenvalues.



[Example of Gershgorin's discs from last recitation.]

Play around with and visualize Gershgorin's theorem and eigenvalues using a Java applet at http://bwlewis.github.io/cassini/.

Sturm sequence

See ref. [10, 11] for proof and explanation of Sturm's sequence property.

See example 5.7 in book for help regarding Q3 in HW3. Also see sec. 4.6.2 in [12].

Also this -
https://www.win.tue.nl/casa/meetings/seminar/previous/_abstract051109_files/presentation_full.pdf

QR method

- Basic QR method - $O(n^3)$
- Hessenberg QR - $O(n^2)$
  - (i) Convert A to tridiagonal/Hessenberg form: $O(n^2)$ steps

- Apply basic QR method
- See [5, 6] for an excellent discussion.
- Watch this (https://www.youtube.com/watch?v=QOfyujCmLGY) for an interesting depiction of how QR algorithm reaches the solution (using Gershgorin's theorem)!

Inverse Iteration Method

- Related to Power method [11].
- Jacobi method can give us eigenvectors - QR/Sturm sequence don't
- Inverse iteration gives eigenvalues and vectors

Let ν be an approximation of an eigenvalue, and $\nu^{(0)}$ the corresponding **approximation** to the eigenvector. Then using the inverse iteration method:

$$
\begin{aligned}
(A - \vartheta I)\boldsymbol{w}^{(k)} &= \boldsymbol{v}^{(k)}, \\
\boldsymbol{v}^{(k+1)} &= c_k \boldsymbol{w}^{(k)},
\end{aligned}
$$

where $c_k = 1/\|w^{(k)}\|_2$, the sequence $\{\nu^{(k)}\}$ converges to the normalized eigenvector $\overline{\nu}$ for the eigenvalue λ closest to ν .

[For proof, see Th. 5.10 in book.]

Computing $w^{(k)}$ at every step requires solving a linear system of equations. Here, we can:

- Use LU decomposition of A
- Convert A to tridiagonal T using Householder's (more efficient)

Rayleigh coefficient

$R(x) = \frac{x^T A x}{x^T x}$ where A is symmetric.

- If x is eigenvector, R(x) is corresponding eigenvalue
- Otherwise, if

$$\boldsymbol{x} = \sum_{j=1}^{n} \alpha_j \, \boldsymbol{x}^{(j)} \,,$$

$$R(\boldsymbol{x}) = \frac{\sum_{j=1}^{n} \lambda_j \, \alpha_j^2}{\sum_{j=1}^{n} \alpha_j^2}$$

- Th. 5.12: $\lambda_{min} \leq R(x) \leq \lambda_{max}$
- If we have a fairly close approximation of the eigenvector, R(x) gives us a good approximate of the corresponding eigenvalue!

Perturbation Analysis

The following theorems (see p. 172 of textbook) give an idea of how a matrix's eigenvalues are affected under perturbation.

Theorem 5.14 *Let $M \in \mathbb{R}^{n\times n}_{\text{sym}}$, with eigenvalues λ_i and corresponding orthonormal eigenvectors $\boldsymbol{v}_i, i = 1, 2, \ldots, n$, and suppose that $\boldsymbol{u} \neq \boldsymbol{0}$ and $\boldsymbol{w}$ are vectors in $\mathbb{R}^n$ and μ is a real number such that*

$$(M - \mu I)\boldsymbol{u} = \boldsymbol{w} \,. \tag{5.44}$$

Then, at least one eigenvalue λ_j of M satisfies

$$|\lambda_j - \mu| \leq \|\boldsymbol{w}\|_2 / \|\boldsymbol{u}\|_2 \,.$$

Theorem 5.15 (Bauer–Fike Theorem (symmetric case)) *Suppose that $A, E \in \mathbb{R}^{n\times n}_{\text{sym}}$ and $B = A - E$. Assume, further, that the eigenvalues of A are denoted by $\lambda_j, j = 1, 2, \ldots, n$, and μ is an eigenvalue of B. Then, at least one eigenvalue λ_j of A satisfies*

$$|\lambda_j - \mu| \leq \|E\|_2 \,.$$

For a general matrix, this result can be extended using the condition number of the diagonalizing/similarity transformation matrix:

$|\lambda_j - \mu| \le k(X)\|E\|_2$ where $X^{-1}AX = D$, and $k(X) = \|X^{-1}\|_2\|X\|_2$ is the condition number of X.

eg.

A = [1 ½ ⅓ ;
½ ⅓ ¼ ;
⅓ ¼ ⅕]

B = [1 0.5 0.3333;
0.5 0.3333 0.25;
0.3333 0.25 0.2]

B is a perturbation (approximation) of A.

$|\lambda_j - \mu| \le \|E\|_2 = \| B - A\|_2 = 3.3 \times 10^{-5}$

In general, symmetric matrices are well-conditioned (robust to perturbation).

Aside about matrix condition numbers

- Relative condition number
  - $sup x(\|f\|/\|f(x)\|)/(\|x\|/\|x\|)$
  - For a matrix (image from Trefethen et al., Numerical Linear Algebra, p. 93)

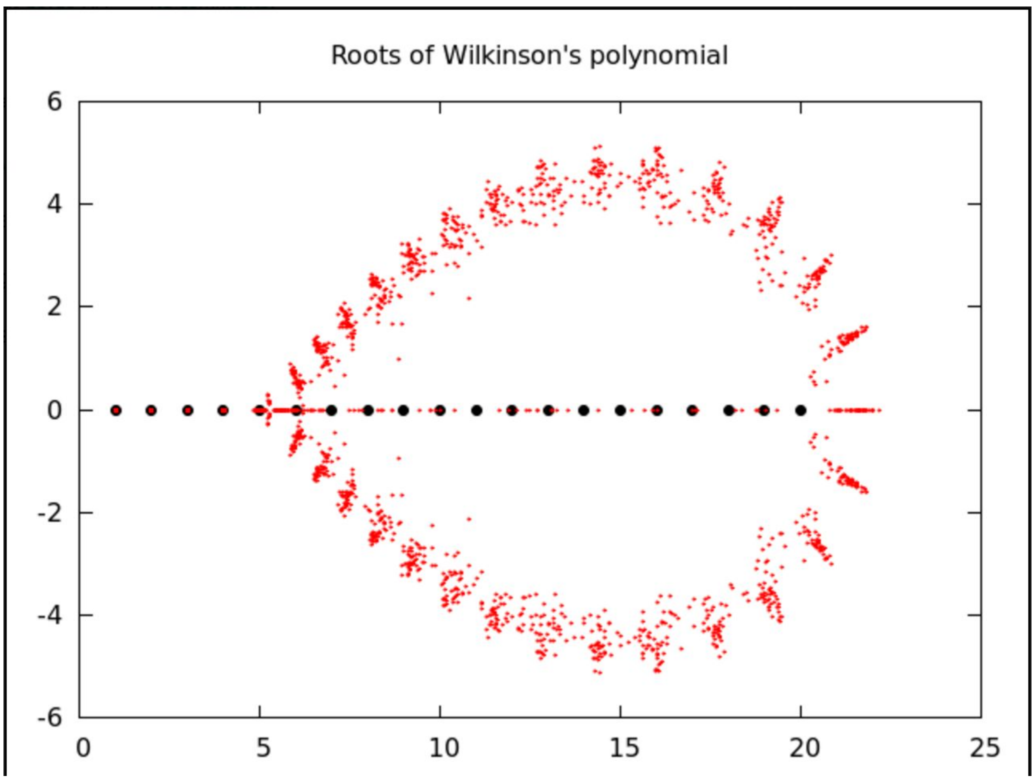
$$\kappa = \sup_{\delta x}\left(\frac{\|A(x+\delta x) - Ax\|}{\|Ax\|} \bigg/ \frac{\|\delta x\|}{\|x\|}\right) = \sup_{\delta x}\frac{\|A\delta x\|}{\|\delta x\|} \bigg/ \frac{\|Ax\|}{\|x\|}$$

that is,

$$\kappa = \|A\|\frac{\|x\|}{\|Ax\|}$$

- Relative condition number more important in numerical analysis, as floating point system introduces relative errors
- Small condition number means well-conditioned. Large means ill-conditioned.
- Condition number of A:
  - $k(A) = \|A\|.\|A^{-1}\|$
- Eg.

- Finding the eigenvalues of a matrix
    - Take A = [1 1000; 0 1], eig(A) = [1 1]
    - B = [1 1000; 0.001 1], eig(B) = [0 2]
    - Problem is ill-conditioned for non-symmetric matrices, well-conditioned for symmetric matrices
- Determining roots of a polynomial is ill-conditioned. If you. See [9] for famous example on Wilkinson's polynomial. You can also try checking this in MATLAB (see [10] for instructions).



- See [8] for more info. on examples.

Helpful links

1. Nice reference on Gerschgorin's theorems
2. Sturm’s theorem
3. Proof of Sturm’s sequence property
4. Using Sturm’s theorem for finding eigenvalues (sec. 4.6.2)
5. QR algorithm explained quite well
6. QR algorithm proof
7. Rayleight coefficient and Inverse iteration
8. *Numerical Linear Algebra, Trefethen & Bau* also has a nice discussion on condition numbers (see Lec. 12).
9. https://en.wikipedia.org/wiki/Wilkinson%27s_polynomial
10. http://blogs.mathworks.com/cleve/2013/03/04/wilkinsons-polynomials/
11. http://college.cengage.com/mathematics/larson/elementary_linear/5e/students/ch08-10/chap_10_3.pdf