Recitation 5

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

Brief Overview

- Underdetermined linear system
- Eigenvalues and eigenvectors
- Jacobi's method for finding eigenvalues
- Application of eigenvalue problem
- Gershgorin's disc theorems
- Examples of the disc theorems using MATLAB (see [7], [8])
- Sturm's sequence property

Underdetermined systems of linear equations

This is a system with fewer equations than unknowns.

Example.

Can you find the solutions for these?

1. $x + y + z = 1$, $x + y + z = 0$
2. $x + y + z = 1$, $x + y + 2z = 3$

Eigenvalue Problems

Theorem 5.1 *Suppose that $A \in \mathbb{R}^{n\times n}_{\text{sym}}$; then, the following statements are valid.*

(i) *There exist n linearly independent eigenvectors $\boldsymbol{x}^{(i)} \in \mathbb{R}^n$ and corresponding eigenvalues $\lambda_i \in \mathbb{R}$ such that $A\boldsymbol{x}^{(i)} = \lambda_i \boldsymbol{x}^{(i)}$ for all $i = 1, 2, \ldots, n$.*

(ii) *The function*

$$\lambda \mapsto \det(A - \lambda I) \tag{5.2}$$

is a polynomial of degree n with leading term $(-1)^n\lambda^n$, called the **characteristic polynomial of** *A. The eigenvalues of A are the zeros of the characteristic polynomial.*

(iii) *If the eigenvalues λ_i and λ_j of A are distinct, then the corresponding eigenvectors $\boldsymbol{x}^{(i)}$ and $\boldsymbol{x}^{(j)}$ are orthogonal in $\mathbb{R}^n$, i.e.,*

$$\boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{x}^{(j)} = 0 \qquad \text{if } \lambda_i \neq \lambda_j\,, \qquad i, j \in \{1, 2, \ldots, n\}\,.$$

(iv) *If λ_i is a root of multiplicity m of (5.2), then there is a linear subspace in $\mathbb{R}^n$ of dimension m, spanned by m mutually orthogonal eigenvectors associated with the eigenvalue λ_i.*

(v) *Suppose that each of the eigenvectors $\boldsymbol{x}^{(i)}$ of A is* **normalised**, *in other words, $\boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{x}^{(i)} = 1$ for $i = 1, 2, \ldots, n$, and let X denote the square matrix whose columns are the normalised (orthogonal) eigenvectors; then, the matrix $\Lambda = X^{\mathrm{T}}AX$ is diagonal, and the diagonal elements of Λ are the eigenvalues of A.*

(vi) *Let $Q \in \mathbb{R}^{n\times n}$ be an orthogonal matrix and define $B \in \mathbb{R}^{n\times n}_{\text{sym}}$ by $B = Q^{\mathrm{T}}AQ$; then, $\det(B - \lambda I) = \det(A - \lambda I)$ for each $\lambda \in \mathbb{R}$. The eigenvalues of B are the same as the eigenvalues of A, and the eigenvectors of B are the vectors $Q^{\mathrm{T}}\boldsymbol{x}^{(i)}$, $i = 1, 2, \ldots, n$.*

(vii) *Any vector $\boldsymbol{v} \in \mathbb{R}^n$ can be expressed as a linear combination of the (ortho)normalised eigenvectors $\boldsymbol{x}^{(i)}$, $i = 1, 2, \ldots, n$, of A, i.e.,*

$$\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}^{(i)}, \qquad \alpha_i = \boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{v}\,.$$

(viii) *The trace of A, $\mathrm{Trace}(A) = \sum_{i=1}^{n} a_{ii}$, is equal to the sum of the eigenvalues of A.*

[Image source: Chap. 5, Introduction to Numerical Analysis, E. Suli & D. Mayers]

- Recap of some important properties (see image from book)
- If $A \in R_{symm}^{n \times n}$ (A is a real, symmetric matrix):
    - There exist 'n' linearly independent eigenvectors
    - $det(\lambda I - A) = 0$ gives the characteristic polynomial (in general)
    - If two eigenvalues are distinct, corresponding eigenvectors are orthogonal
    - Other properties in textbook.
- **Why are Eigenvalue problems important?** Ubiquitous in numerical linear algebra, especially solving ODEs, modelling symmetric physical systems or laws etc.
- **How can we calculate Eigenvalues of large matrices?**
    - **method 1**: Write down characteristic polynomial, and find its roots numerically. This is not very practical for three reasons:
        i. A 100x100 matrix will have 100 eigenvalues. Newton's method works well when starting very close to the optimal value, but may diverge otherwise.
        ii. May divide the characteristic polynomial once a root has been found, but polynomial division can be numerically dangerous/unstable.
        iii. To find eigenvectors, still have to solve 'n' linear equations, which will take $O(n^3)$ time!
        iv. Note: In general, **polynomial root-finding is an ill-conditioned problem**. See eg. 5.12 (p. 92, Numerical Linear Algebra, Trefethen & Bau) on Wilkinson's polynomial.
    - **method 2**: Use an iterative solving method (like Newton's method) which diagonalizes the matrix! Some popular methods are:
        i. Jacobi
        ii. QR
        iii. Sturm sequence
        iv. Power (and power-inverse)

- A bad example. Consider the matrix:

$$A = \begin{pmatrix} 0 & & & & & & \varepsilon \\ 1 & 0 & & & & & \\ & 1 & 0 & & & & \\ & & 1 & 0 & & & \\ & & & 1 & 0 & & \\ & & & & \ddots & \ddots & \\ & & & & & 1 & 0 \end{pmatrix}$$

  - Charac. polynomial: $\lambda^n - \varepsilon = 0$
    i. Case 1: $\varepsilon = 0, \lambda_i = 0$
    ii. Case 2: $\varepsilon = 10^{-40}$, relative error = $10^{-40}/1 = 10^{-40}$. One eigenvalue, $\lambda_k = 1/10 = 0.1$. Thus, adding an epsilon term changes one eigenvalue by $10^{39} \times \varepsilon$ times! This is an ill-conditioned problem, and numerically unstable.
  - See [1] for more info. on this problem.

Jacobi's method

Idea: Use orthogonal transformations (pre- and post- multiply) to convert matrix to diagonal form.

- Use a plane rotation matrix of the form:

$$R(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

- Can check that this is an orthogonal matrix.
- Choose ϕ to make (p,q) and (q,p) element zero:

$$\varphi = \frac{1}{2}\tan^{-1}\frac{2a_{pq}}{a_{qq} - a_{pp}}$$

- See [2] for example. Also see sec. 5.2 in [1] for good discussion.

Gershgorin's theorems

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.

Visualizing Gershgorin's discs

Examples:

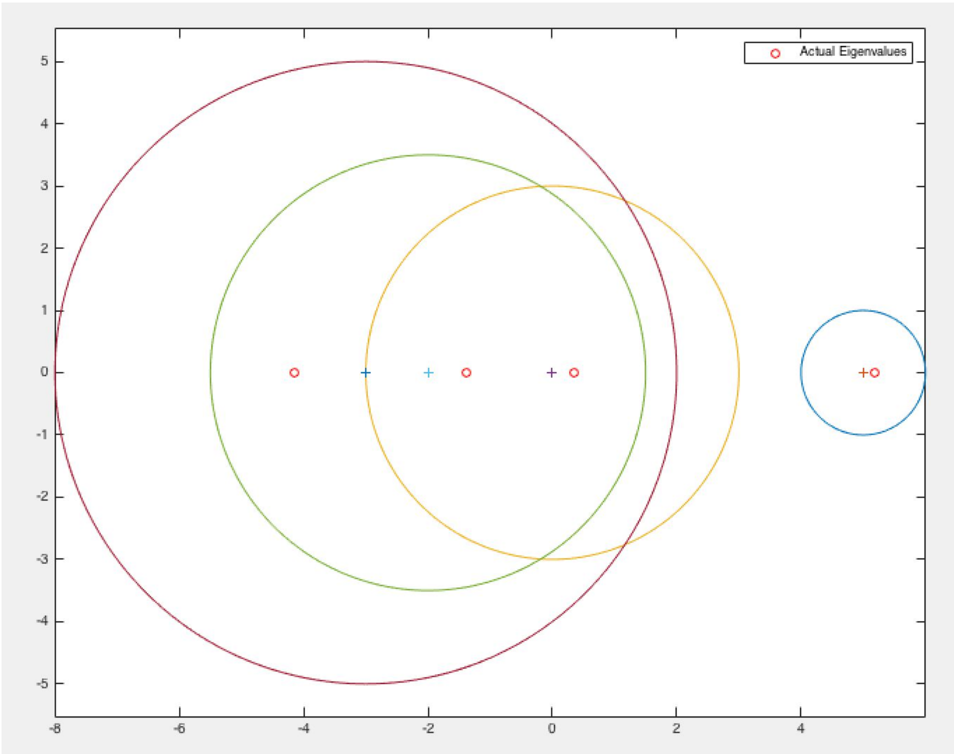
1. A = [1 2; 1 -1]
2. A = [1 -1; 2 -1]
3. A = [5 0 0 -1; 1 0 -1 1; -1.5 1 -2 1; -1 1 3 -3]

You can visualize the Gershgorin discs for the above examples in MATLAB as follows:

1. Define a function gershdisc.m using the code below.
2. Define A = [...] from the example above.
3. Run `gershdisc(A)` in the folder where you saved this function.
4. Try to understand what the function is doing. Try some examples of your own!

```
% gershdisc.m
% This function plots the Gershgorin Discs for the matrix A passed as an argument.
% It will also plot the centers of such discs, and the actual eigenvalues
% of the matrix.
 function gershdisc(A)

error(nargchk(nargin,1,1));
if size(A,1) ~= size(A,2)
   error('Matrix should be square');
   return;
end

for i=1:size(A,1)
   h=real(A(i,i)); k=imag(A(i,i));
   r=0;
   for j=1:size(A,1)
      if i ~= j
         r=r+(norm(A(i,j)));
      end
   end
   N=256;
   t=(0:N)*2*pi/N;
   plot( r*cos(t)+h, r*sin(t)+k ,'-');
   hold on;
   plot( h, k,'+');
end
axis equal;
ev=eig(A);
for i=1:size(ev)
   rev=plot(real(ev(i)),imag(ev(i)),'ro');
end
hold off;
legend(rev,'Actual Eigenvalues');

end

%code source: http://www.mathworks.com/matlabcentral/fileexchange/13989-gershgorin-discs-plot/content/gershdisc.m
```

For eg. 3, you should get:



[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

[HW3 tips]

For Q3, you can use a function SturmSequence(A) which returns the eigen values of A. You can use a separate function for finding out $p_i(v)$, which takes in (A,i,v) as parameters for a matrix A and submatrix 'i' and returns $det(T_i - vI)$, which is just $p_i(v)$.

Helpful links

1. See p. 1 & 2 on bad eigen value problems
2. See sec. 2 on Jacobi method
3. More about Jacobi's method
4. Nice reference on Gerschgorin's theorems
5. Found this nice MATLAB tutorial
6. Quick overview of linear algebra and relevant numerical algorithms
7. Jacobi convergence and eigenvalue problem examples
8. See topics ‘markov chain 1 / 2’ for applications of eigenvalue problem in probability
9. Underdetermined systems
10. Sturm’s theorem
11. Proof of Sturm’s sequence property
12. Using Sturm’s theorem for finding eigenvalues (sec. 4.6.2)