MA 8019: Numerical Analysis I
Eigenvalue Problems

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan
E-mail: syyang@math.ncu.edu.tw
http://www.math.ncu.edu.tw/∼syyang/

First version: May 05, 2018   Last updated: December 2, 2023

Vector space $\mathbb{C}^n$

- **Field $\mathbb{C}$**

  $\mathbb{C}$ is the set of complex numbers have the form $\gamma = \alpha + i\beta$, where $\alpha, \beta \in \mathbb{R}$ and $i = \sqrt{-1}$.

  The conjugate of γ is defined by $\bar{\gamma} := \alpha - i\beta$. Then $\bar{\bar{\gamma}} = \gamma$.

  The modulus of γ is defined by $|\gamma| := \sqrt{\alpha^2 + \beta^2}$. Then $\gamma\bar{\gamma} = (\alpha + i\beta)(\alpha - i\beta) = \alpha^2 - i\alpha\beta + i\alpha\beta - i^2\beta^2 = \alpha^2 + \beta^2 = |\gamma|^2$.

- **Vector space $\mathbb{C}^n$**

  $\mathbb{C}^n$ denotes the set of all complex n-tuples, $x = (x_1, x_2, \ldots, x_n)^\top$, $x_j \in \mathbb{C}$ for $1 \le j \le n$, such that for any $x, y \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ we define

$$
\begin{aligned}
x + y &:= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^\top, \\
\lambda x &:= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)^\top.
\end{aligned}
$$

Inner product, Euclidean norm, and conjugate transpose

- **Inner product and Euclidean norm**

  For $x, y \in \mathbb{C}^n$,

$$\langle x, y\rangle := \sum_{j=1}^{n} x_j \bar{y}_j, \quad \|x\|_2 := \sqrt{\langle x, x\rangle} = \sqrt{\sum_{j=1}^{n} x_j \bar{x}_j} = \sqrt{\sum_{j=1}^{n} |x_j|^2}.$$

- **Conjugate transpose**

  Let $A \in \mathbb{C}^{m\times n}$ be a matrix having complex entries. The conjugate transpose of A is defined as $A^* := \bar{A}^\top$, that is, $(A^*)_{jk} = \bar{A}_{kj}$.

  If $x = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{C}^{n\times 1} = \mathbb{C}^n$ is an $n \times 1$ matrix (or column vector), then $x^* = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) \in \mathbb{C}^{1\times n}$ is a $1 \times n$ matrix (or row vector), and

$$\langle x, y\rangle = y^* x, \quad \|x\|_2^2 = \langle x, x\rangle = x^* x.$$

Eigenvalue and eigenvector

- **Definition:** Let $A \in \mathbb{C}^{n\times n}$. Then $\lambda \in \mathbb{C}$ is an eigenvalue of A if $\exists\, 0 \neq x \in \mathbb{C}^n$ (called eigenvector) such that

$$Ax = \lambda x.$$

- **Example:**

$$\begin{bmatrix} 2 & 0 & 1 \\ 5 & -1 & 2 \\ -3 & 2 & -\frac{5}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} = (-2) \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}.$$

-2 is an eigenvalue of A,
$(1, 3, -4)^\top$ is a corresponding eigenvector.

- **Equivalent statements:**

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A \quad &\Longleftrightarrow \quad A - \lambda I \text{ is singular} \\ &\Longleftrightarrow \quad p(\lambda) := \det(A - \lambda I) = 0. \end{aligned}$$

The polynomial $p(\lambda)$ in λ is called the characteristic polynomial of A.

Some remarks

- The direct method may be the best method for computing eigenvalues for small size matrices.
- For large size matrices, we have to find the roots of the characteristic polynomial of high degree, which is an *ill-conditioned problem* in general. That is, small changes in the coefficients of polynomial yield huge changes in the roots.

  See Example 2 in Chapter 2 on page 68.

Gershgorin's Theorem

Gershgorin's Theorem: *The spectrum of an $n \times n$ matrix (that is, the set of its eigenvalues) is contained in the union of the following n disks D_i $(1 \le i \le n)$ in the complex plane:*

$$D_i = \{z \in \mathbb{C} : |z - a_{ii}| \le \sum_{j=1, j\neq i}^{n} |a_{ij}|\} \quad (1 \le i \le n).$$

Proof: Let λ be an eigenvalue of A. Then there exists $x \neq 0$ with $\|x\|_\infty = 1$ such that $Ax = \lambda x$. Let i be an index for which $\|x\|_\infty = |x_i| = 1$. Since $(Ax)_i = \lambda x_i$, we have

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j \quad \Longrightarrow \quad (\lambda - a_{ii}) x_i = \sum_{j=1, j\neq i}^{n} a_{ij} x_j.$$

Therefore,

$$|\lambda - a_{ii}| = |\lambda - a_{ii}||x_i| \le \sum_{j=1, j\neq i}^{n} |a_{ij}||x_j| \le \sum_{j=1, j\neq i}^{n} |a_{ij}|.$$

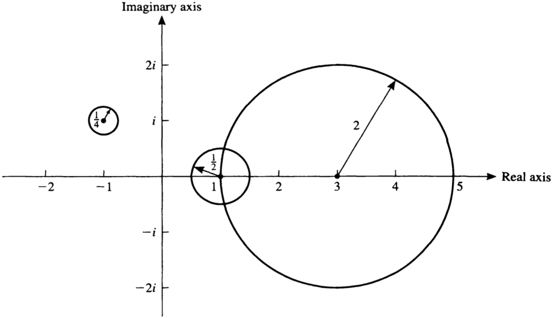
Thus, $\lambda \in D_i$.

Example

Locate the eigenvalues of the matrix A without computing them:

$$A = \begin{bmatrix} -1+i & 0 & \frac{1}{4} \\ \frac{1}{4} & 1 & \frac{1}{4} \\ 1 & 1 & 3 \end{bmatrix}.$$

All eigenvalues of A satisfy the inequality $\frac{1}{2} \le |\lambda| \le 5$.



Power method: assumptions

Power method is designed to compute the largest eigenvalue and corresponding eigenvector of a matrix $A \in \mathbb{C}^{n\times n}$ satisfying the following two assumptions:

- There is a single eigenvalue of maximum modulus, i.e.,

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_{n-1}| \geq |\lambda_n|.$$

- There is a linear independent set of n eigenvectors, i.e.,

  $\exists \{u^{(1)}, u^{(2)}, \cdots, u^{(n)}\}$ such that $Au^{(j)} = \lambda_j u^{(j)}, 1 \leq j \leq n,$

  and $\{u^{(1)}, u^{(2)}, \cdots, u^{(n)}\}$ is a basis of $\mathbb{C}^n$.

Power method

- Let $x^{(0)}$ be a nonzero vector in $\mathbb{C}^n$. Assume that

$$x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \cdots + a_n u^{(n)},$$

  and $a_1 \neq 0$.

- Define $x^{(k)} = Ax^{(k-1)}$, so that

$$\begin{aligned} x^{(k)} &= A^k x^{(0)} = a_1 A^k u^{(1)} + a_2 A^k u^{(2)} + \cdots + a_n A^k u^{(n)} \\ &= a_1 \lambda_1^k u^{(1)} + a_2 \lambda_2^k u^{(2)} + \cdots + a_n \lambda_n^k u^{(n)} \\ &= \lambda_1^k \left\{ a_1 u^{(1)} + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k u^{(2)} + \cdots + a_n \left(\frac{\lambda_n}{\lambda_1}\right)^k u^{(n)} \right\}. \end{aligned}$$

- Since $|\lambda_1| > |\lambda_j|$ for $2 \leq j \leq n$, then $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$ as $k \to \infty$.

  Therefore, we may assume that $x^{(k)} = \lambda_1^k (a_1 u^{(1)} + \varepsilon^{(k)})$, where $\varepsilon^{(k)} \to 0$ as $k \to \infty$.

Power method (cont'd)

- Let $\Phi : \mathbb{C}^n \to \mathbb{C}$ be a linear functional (linear transformation), i.e., $\Phi(\alpha x + \beta y) = \alpha\Phi(x) + \beta\Phi(y), \forall\, x, y \in \mathbb{C}^n, \alpha, \beta \in \mathbb{C}$,
  Then Φ is a continuous function on the normed vector space $(\mathbb{C}^n, \|\cdot\|_2)$ and $\Phi(0) = 0$.

  Example: $\Phi : \mathbb{C}^n \to \mathbb{C}$, $\Phi((x_1, x_2, \cdots, x_n)^\top) = x_j$, i.e., Φ chooses the j-th component of a vector.

- Apply Φ to $x^{(k)} = \lambda_1^k(a_1 u^{(1)} + \varepsilon^{(k)})$. Then

$$\begin{aligned}\Phi(x^{(k)}) &= \lambda_1^k \left\{a_1\Phi(u^{(1)}) + \Phi(\varepsilon^{(k)})\right\}, \\ \Phi(x^{(k+1)}) &= \lambda_1^{k+1} \left\{a_1\Phi(u^{(1)}) + \Phi(\varepsilon^{(k+1)})\right\}.\end{aligned}$$

- Define $r_k := \dfrac{\Phi(x^{(k+1)})}{\Phi(x^{(k)})}$. Then

$$r_k = \lambda_1 \left\{ \frac{a_1\Phi(u^{(1)}) + \Phi(\varepsilon^{(k+1)})}{a_1\Phi(u^{(1)}) + \Phi(\varepsilon^{(k)})} \right\} \to \lambda_1 \text{ as } k \to \infty.$$

Algorithm for the power method

input n, A, x, m
output r, x
for $k = 1$ **to** m **do**
    $y \leftarrow Ax$
    $r \leftarrow \Phi(y)/\Phi(x)$
    $x \leftarrow y/\|y\|$ (In general, we take $\|\cdot\| = \|\cdot\|_\infty$)
    output k, x, r
end do

Note:

- Here Φ is some linear functional.
- *We normalize $x^{(k)}$ at each step to avoid $\|x^{(k)}\| \to 0$ or ∞. The ratios r are the same as those in the unnormalized version (page 262, # 2).*

Example

Use the power method on the following matrix and initial vector:

$$A = \begin{bmatrix} 6 & 5 & -5 \\ 2 & 6 & -2 \\ 2 & 5 & -1 \end{bmatrix}, \quad x^{(0)} = (-1,1,1)^\top, \quad \text{eig}(A) = \{6,4,1\}.$$

We take $\Phi(x) = x_2$. The numerical results are shown below:

$$\begin{array}{lll}
k=0 & x^{(0)} = (-1.00000, \quad 1.00000, \quad 1.00000) & \\
k=1 & x^{(1)} = (-1.00000, \quad 0.33333, \quad 0.33333) & r_0 = \quad 2.0 \\
k=2 & x^{(2)} = (-1.00000, -0.11111, -0.11111) & r_1 = -2.0 \\
k=3 & x^{(3)} = (-1.00000, -0.40741, -0.40741) & r_2 = \ 22.0 \\
k=4 & x^{(4)} = (-1.00000, -0.60494, -0.60494) & r_3 = \quad 8.9091 \\
\vdots & \quad \vdots & \quad \vdots \\
k=6 & x^{(6)} = (-1.00000, -0.82442, -0.82442) & r_5 = \quad 6.71508 \\
\vdots & \quad \vdots & \quad \vdots \\
k=28 & x^{(28)} = (-1.00000, -0.99998, -0.99998) & r_{27} = \quad 6.00007
\end{array}$$

leading eigenvalue: 6 corresponding eigenvector: $(-1,-1,-1)^\top$

Inverse power method

Note: *Assume that $A \in \mathbb{C}^{n\times n}$ is a nonsingular matrix. If λ is an eigenvalue of A, then λ^{-1} is an eigenvalue of A^{-1}.*
Proof: $\exists\, x \neq 0$ s.t. $Ax = \lambda x \Longrightarrow x = \lambda A^{-1}x \Longrightarrow A^{-1}x = \lambda^{-1}x$.

Suppose that the eigenvalues of A can be arranged as follows:

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{n-1}| > |\lambda_n| > 0.$$

Then the eigenvalues of A^{-1} are given by

$$|\lambda_n^{-1}| > |\lambda_{n-1}^{-1}| \geq \cdots \geq |\lambda_2^{-1}| \geq |\lambda_1^{-1}| > 0.$$

We can compute the smallest eigenvalue λ_n of A by applying the power method to A^{-1},

$$x^{(k+1)} = A^{-1}x^{(k)} \Longrightarrow \text{we solve } \; Ax^{(k+1)} = x^{(k)},$$

where we can carry out the LU decomposition only once.

Example

First, we can find the LU decomposition of A by GE:

$$A = \begin{bmatrix} 6 & 5 & -5 \\ 2 & 6 & -2 \\ 2 & 5 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{3} & \frac{10}{13} & 1 \end{bmatrix} \begin{bmatrix} 6 & 5 & -5 \\ 0 & \frac{13}{3} & \frac{-1}{3} \\ 0 & 0 & \frac{12}{13} \end{bmatrix}.$$

We begin with the initial vector $x^{(0)} = (3, 7, -13)^\top$. The results of the inverse power method are given below:

$$\begin{array}{lll}
k = 0 & x^{(0)} = (\ \ 3.00000,\ \ 7.00000, -13.00000) & \\
k = 1 & x^{(1)} = (-0.80165, -0.00826, \ -1.00000) & r_0 = -5.8889 \\
k = 2 & x^{(2)} = (-0.95089, -0.01774, \ -1.00000) & r_1 = \ \ 1.19759 \\
k = 3 & x^{(3)} = (-0.98759, -0.00712, \ -1.00000) & r_2 = \ \ 1.02750 \\
k = 4 & x^{(4)} = (-0.99688, -0.00223, \ -1.00000) & r_3 = \ \ 1.00446 \\
\vdots & \vdots & \vdots \\
k = 6 & x^{(6)} = (-0.99980, -0.00017, \ -1.00000) & r_5 = \ \ 1.00012 \\
\vdots & \vdots & \vdots \\
k = 11 & x^{(11)} = (-1.00000, \ \ 0.00000, \ -1.00000) & r_{10} = \ \ 1.00000
\end{array}$$

smallest eigenvalue of A: 1 corresponding eigenvector: $(-1, 0, -1)^\top$

Shifted inverse power method

- The power method computes the largest eigenvalue of A. The inverse power method computes the smallest eigenvalue of A.

  Let $\mu \in \mathbb{C}$ be a given number. How to compute the eigenvalue of A closest to the given μ?

- Suppose that eigenvalue λ_k of A satisfies $0 < |\lambda_k - \mu| < \varepsilon$, and all other eigenvalues of A satisfy $|\lambda_j - \mu| > \varepsilon$ for some $\varepsilon > 0$.

  Since the eigenvalues of $A - \mu I$ are $\lambda_j - \mu$, we can apply the inverse power method on $A - \mu I$ (power method to $(A - \mu I)^{-1}$) to compute $z = (\lambda_k - \mu)^{-1}$. Then $\lambda_k = z^{-1} + \mu$.

  This method is called the shifted inverse power method.