Numerical Analysis

Lecture 6: Approximating Eigenvalues

Instructor: Prof. Guanding Yu
Zhejiang University

Outline

Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

Eigenvalues and Eigenvectors

We have

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

Definition (Characteristic Polynomial)

If $\mathbf{A}$ is a square matrix, the characteristic polynomial of $\mathbf{A}$ is defined by

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$$

Eigenvalues and Eigenvectors

Definition (Characteristic Polynomial)

If $\mathbf{A}$ is a square matrix, the characteristic polynomial of $\mathbf{A}$ is defined by

$$p(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I})$$

Comments

- p is an nth-degree polynomial and, consequently, has at most n distinct zeros, some of which might be complex.
- If λ is a zero of p, then, since $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$, we can prove that the linear system defined by

  $$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$$

  has a solution with $\mathbf{x} \neq \mathbf{0}$.

Eigenvalues and Eigenvectors

Finding the Eigenvalues & Eigenvectors

- To determine the eigenvalues of a matrix, we can use the fact that λ is an eigenvalue of A if and only if

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

- Once an eigenvalue λ has been found, a corresponding eigenvector $\mathbf{x} \neq \mathbf{0}$ is determined by solving the system

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

Eigenvalues and Eigenvectors

Difficulties

- Finding the determinant of an $n \times n$ matrix is computationally expensive.
- Finding good approximations to the roots of $\rho(\lambda)$ is also difficult.
- The computational complexity of solving a linear system is high.

Linear Algebra

Linearly Independent

Let $\{\mathbf{v}^{(\mathbf{1})}, \mathbf{v}^{(\mathbf{2})}, \cdots, \mathbf{v}^{(\mathbf{k})}\}$ be a set of vectors. The set is linearly independent if whenever

$$\mathbf{0} = \alpha_{\mathbf{1}}\mathbf{v}^{(\mathbf{1})} + \alpha_{\mathbf{2}}\mathbf{v}^{(\mathbf{2})} + \cdots + \alpha_{\mathbf{k}}\mathbf{v}^{(\mathbf{k})},$$

then $\alpha_i = 0, \forall i$. Otherwise the set of vectors is linearly dependent.

Linearly Combination and Basis

Suppose that $\{\mathbf{v}^{(\mathbf{1})}, \mathbf{v}^{(\mathbf{2})}, \cdots, \mathbf{v}^{(\mathbf{k})}\}$ is a set of k linearly independent vectors. Then, for any vector $\mathbf{x}$, a unique collection of constants $\beta_1, \beta_2, \cdots, \beta_k$ exists with

$$\mathbf{x} = \beta_{\mathbf{1}}\mathbf{v}^{(\mathbf{1})} + \beta_{\mathbf{2}}\mathbf{v}^{(\mathbf{2})} + \cdots + \beta_{\mathbf{k}}\mathbf{v}^{(\mathbf{k})}.$$

Linear Algebra

Linearly Independent Eigenvectors

If A is a matrix and $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of A with associated eigenvectors $\mathbf{x}^{(\mathbf{1})}, \mathbf{x}^{(\mathbf{2})}, \cdots, \mathbf{x}^{(\mathbf{k})}$, then $\{\mathbf{x}^{(\mathbf{1})}, \mathbf{x}^{(\mathbf{2})}, \cdots, \mathbf{x}^{(\mathbf{k})}\}$ is a linearly independent set.

Orthogonal Vectors

A set of vectors $\left\{v^{(1)}, v^{(2)}, \cdots, v^{(n)}\right\}$ is called orthogonal if $(v^{(i)})^t v^{(j)} = 0$, for all $i \neq j$. If, in addition, $(v^{(i)})^t v^{(i)} = 1$, for all $i = 1, 2, \cdots, n$, then the set is called orthonormal.

Orthogonal Matrices

Definition

A matrix Q is said to be orthogonal if its columns $\{\mathbf{q_1^t}, \mathbf{q_2^t}, \cdots, \mathbf{q_n^t}\}$ form an orthonormal set in $\mathbb{R}^n$.

Properties

Suppose that Q is an orthogonal $n \times n$ matrix. Then

- Q is invertible with $Q^{-1} = Q^t$;
- For any $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, $(Q\mathbf{x})^{\mathbf{t}}\mathbf{Qy} = \mathbf{x^t y}$;
- For any $\mathbf{x}$ in $\mathbb{R}^n$, $\| Q\mathbf{x} \|_2 = \| \mathbf{x} \|_2$.

Similar Matrices

Definition

Two matrices A and B are said to be similar if a nonsingular matrix S exists with $A = S^{-1}BS$.

Properties

Suppose A and B are similar matrices with $A = S^{-1}BS$ and λ is an eigenvalue of A with associated eigenvector $\mathbf{x}$. Then λ is an eigenvalue of B with associated eigenvector $\mathbf{Sx}$.

Diagonally Similar Matrices

Definition

An $n \times n$ matrix A is similar to a diagonal matrix D if and only if A has n linearly independent eigenvectors. In this case, $D = S^{-1}AS$, where the columns of S consist of the eigenvectors, and the ith diagonal element of D is the eigenvalue of A that corresponds to the ith column of S.

Properties

An $n \times n$ matrix A that has n distinct eigenvalues is similar to a diagonal matrix.

Symmetric Matrices

Definition

An $n \times n$ matrix A is symmetric if and only if there exists a diagonal matrix D and an orthogonal matrix Q with $A = QDQ^t$.

Properties

- Suppose that A is a symmetric $n \times n$ matrix. There exist n eigenvectors of A that form an orthonormal set, and the eigenvalues of A are real numbers.
- A symmetric matrix A is positive definite if and only if all the eigenvalues of A are positive.

Geršgorin Circle

Let A be an $n \times n$ matrix and R_i denote the circle in the complex plane with center a_{ii} and radius $\sum_{j=1, j\neq i}^{n} |a_{ij}|$; that is,

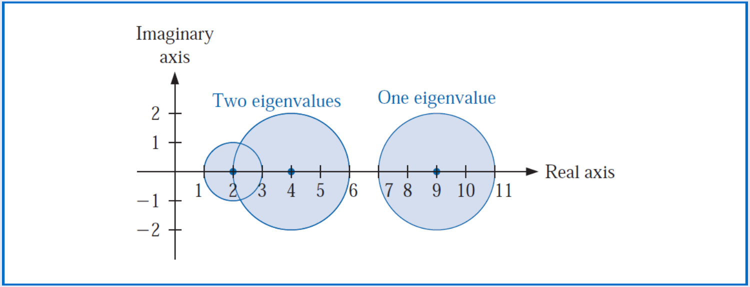
$$R_i = \left\{ z \in \mathcal{C} : |z - a_{ii}| \leq \sum_{j=1, j\neq i}^{n} |a_{ij}| \right\}$$

where $\mathcal{C}$ denotes the complex plane. The eigenvalues of A are contained within the union of theses circles, $R = \bigcup_{i=1}^{n} R_i$. Moreover, the union of any k of the circles that do not intersect the remaining $(n-k)$ contains precisely k (counting multiplicities) of the eigenvalues.

Geršgorin Circle: Examples

Determine the Geršgorin circle for the matrix and use these to find bounds for the spectral radius of A.

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{bmatrix}$$



Outline

The Power Method

Overview

- The Power method is an iterative technique used to determine the dominant eigenvalue of a matrix, that is, the eigenvalue with the largest magnitude.
- One useful feature of the Power method is that it produces not only an eigenvalue, but also an associated eigenvector.
- The Power method is often applied to find an eigenvector for an eigenvalue that is determined by some other means.

The Power Method

Overview

Assume an $n \times n$ matrix A has precisely one eigenvalue, λ_1, that is largest in magnitude, that is

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \geq 0.$$

The Power Method

Observation 1/2

If $\mathbf{x}$ is any vector in $\mathbb{R}^n$, the fact that $\{\mathbf{v^{(1)}}, \mathbf{v^{(2)}}, \cdots, \mathbf{v^{(n)}}\}$ is linearly independent implies that constants $\beta_1, \beta_2, \cdots, \beta_n$ exist with

$$\mathbf{x} = \beta_{\mathbf{1}}\mathbf{v^{(1)}} + \beta_{\mathbf{2}}\mathbf{v^{(2)}} + \cdots + \beta_{\mathbf{n}}\mathbf{v^{(n)}}.$$

Multiplying both sides of this equation by $A, A^2, \cdots, A^k, \cdots$ gives

$$Ax = \sum_{j=1}^{n} \beta_j A v^{(j)} = \sum_{j=1}^{n} \beta_j \lambda_j v^{(j)}$$

$$A^k x = \sum_{j=1}^{n} \beta_j A^k v^{(j)} = \sum_{j=1}^{n} \beta_j \lambda_j^k v^{(j)}$$

The Power Method

Observation 2/2

We further have

$$A^k x = \lambda_1^k \sum_{j=1}^{n} \beta_j \left(\frac{\lambda_j}{\lambda_1} \right)^k v^{(j)},$$

and

$$\lim_{k \to \infty} A^k x = \lim_{k \to \infty} \lambda_1^k \beta_1 v^{(1)}.$$

The Power Method

The Iterative Procedure

Let $X_0 = [1, 1, \cdots, 1]^t$, $Y_k = AX_k$, and $X_{k+1} = \frac{1}{c_{k+1}} Y_k$ where $c_{k+1} = \|Y_k\|_\infty$. Then, we have

$$\lim_{k\to\infty} X_k = v^{(1)}, \lim_{k\to\infty} c_k = \lambda_1.$$

Proof 1/3

We can easily prove that

$$X_k = \frac{\lambda_1^k}{c_1 c_2 \cdots c_k} \sum_{j=1}^{n} \beta_j \left(\frac{\lambda_j}{\lambda_1}\right)^k v^{(j)},$$

The Power Method

Proof 2/3

This means that

$$\lim_{k\to\infty} X_k = \lim_{k\to\infty} \frac{\beta_1 \lambda_1^k}{c_1 c_2 \cdots c_k} v^{(1)}.$$

Because both $\|X_k\|_\infty = \|v^{(1)}\|_\infty = 1$, we have

$$\lim_{k\to\infty} \frac{\beta_1 \lambda_1^k}{c_1 c_2 \cdots c_k} = 1, \textbf{ and } \lim_{k\to\infty} X_k = v^{(1)}.$$

The Power Method

Proof 3/3

Furthermore, from

$$\frac{\beta_1 \lambda_1^k}{c_1 c_2 \cdots c_k} \approx 1,$$

and

$$\frac{\beta_1 \lambda_1^{k-1}}{c_1 c_2 \cdots c_{k-1}} \approx 1,$$

we have $\lambda_1 \approx c_k$.

The matrix

$$A = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix}$$

Has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 1$ with corresponding eigenvectors $\mathbf{v}_1 = (1, -2)^t$ and $\mathbf{v}_2 = (1, -1)^t$. If we start with the arbitrary vector $\mathbf{x}_0 = (1, 1)^t$ and multiply by the matrix A we obtain

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} -5 \\ 13 \end{bmatrix}, \qquad \mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} -29 \\ 61 \end{bmatrix}, \qquad \mathbf{x}_3 = A\mathbf{x}_2 = \begin{bmatrix} -125 \\ 253 \end{bmatrix},$$

$$\mathbf{x}_4 = A\mathbf{x}_3 = \begin{bmatrix} -509 \\ 1021 \end{bmatrix}, \qquad \mathbf{x}_5 = A\mathbf{x}_4 = \begin{bmatrix} -2045 \\ 4093 \end{bmatrix}, \qquad \mathbf{x}_6 = A\mathbf{x}_5 = \begin{bmatrix} -8189 \\ 16381 \end{bmatrix}.$$

As a consequence, approximations to the dominant eigenvalue $\lambda_1 = 4$ are

$$\lambda_1^{(1)} = \frac{61}{13} = 4.6923, \qquad \lambda_1^{(2)} = \frac{253}{61} = 4.14754, \qquad \lambda_1^{(3)} = \frac{1021}{253} = 4.03557,$$

$$\lambda_1^{(4)} = \frac{4093}{1021} = 4.00881, \qquad \lambda_1^{(5)} = \frac{16381}{4093} = 4.00200.$$

An approximate eigenvector corresponding to $\lambda_1^{(5)} = \dfrac{16381}{4093} = 4.00200$ is

$$\mathbf{x}_6 = \begin{bmatrix} -8189 \\ 16381 \end{bmatrix}, \quad \text{which, divided by 16381, normalizes to} \quad \begin{bmatrix} -0.49908 \\ 1 \end{bmatrix} \approx \mathbf{v}_1.$$

The Power Method

Discussions

- The method fails if there is not a single dominant eigenvalues.
- The method fails if the initial vector, X_0 contains a zero contribution from the eigenvector associated with the dominant eigenvalue, i.e., $\beta_1 = 0$.
- How to find the other eigenvalues?

Outline

The Inverse Power Method

Motivation

The Inverse Power method is a modification of the Power method that gives faster convergence. It is used to determine the eigenvalue of A that is closest to a specified number q.

The Inverse Eigenvalues

Suppose the matrix A has eigenvalues $\lambda_1, \cdots, \lambda_n$ with linearly independent eigenvectors $\mathbf{v^{(1)}}, \mathbf{v^{(2)}}, \cdots, \mathbf{v^{(n)}}$. Then, the eigenvalues of $(A - qI)^{-1}$, where $q \neq \lambda_i, \forall i$, are

$$\frac{1}{\lambda_1 - q}, \frac{1}{\lambda_2 - q}, \cdots, \frac{1}{\lambda_n - q},$$

with these same eigenvectors $\mathbf{v^{(1)}}, \mathbf{v^{(2)}}, \cdots, \mathbf{v^{(n)}}$.

Assignment

- Reading Assignment: Chap 9.1, 9.2, 9.3.
- Assignment (optional): Exercise 9.3.1 on Page 590.