CSE 275 Matrix Computation

Ming-Hsuan Yang

Electrical Engineering and Computer Science
University of California at Merced
Merced, CA 95344
http://faculty.ucmerced.edu/mhyang

Lecture 13

Overview

- Eigenvalue problem
- Schur decomposition
- Eigenvalue algorithms

Reading

- Chapter 24 of *Numerical Linear Algebra* by Llyod Trefethen and David Bau
- Chapter 7 of *Matrix Computations* by Gene Golub and Charles Van Loan

Eigenvalues and eigenvectors

- Let $A \in \mathbb{C}^{m \times m}$ be a square matrix, a nonzero $\mathbf{x} \in \mathbb{C}^m$ is an eigenvector of A, and $\lambda \in \mathbb{C}$ is its corresponding eigenvalue if

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

- Idea: the action of a matrix A on a subspace $S \in \mathbb{C}^m$ may sometimes mimic scalar multiplication
- When it happens, the special subspace S is called an eigenspace, and any nonzero $\mathbf{x} \in S$ is an eigenvector
- The set of all eigenvalues of a matrix A is the spectrum of A, a subset of $\mathbb{C}$ denoted by $\Lambda(A)$

Eigenvalues and eigenvectors (cont'd)

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

- Algorithmically: simplify solutions of certain problems by reducing a coupled system to a collection of scalar problems
- Physically: give insight into the behavior of evolving systems governed by linear equations, e.g., resonance (of musical instruments when struck or plucked or bowed), stability (of fluid flows with small perturbations)

Eigendecomposition

- An eigendecomposition (eigenvalue decomposition) of a square matrix A is a factorization

$$A = X \Lambda X^{-1}$$

  where X is a nonsingular and Λ is diagonal

- Equivalently,

$$AX = X\Lambda$$

$$A \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}$$

$$A\mathbf{x}_j = \lambda_j \mathbf{x}_j$$

- λ_j is an eigenvalue and j-th column of X is the corresponding eigenvector

Interpretation using eigendecomposition

- Express a change of basis to "eigenvector coordinates"
- Let $A\mathbf{x} = \mathbf{b}$ and $A = X\Lambda X^{-1}$, we have

$$(X^{-1}\mathbf{b}) = \Lambda(X^{-1}\mathbf{x})$$

- Thus, to compute $A\mathbf{x}$,
  - we can expand $\mathbf{x}$ in the basis of columns of X, apply Λ,
  - and interpret the result as a vector of coefficients of a linear combination of the columns of X

Geometric multiplicity

- The set of eigenvectors corresponding to a single eigenvalue λ, together with the zero vector, forms a subspace of $\mathbb{C}^m$ known as an eigenspace, E_λ
- An eigenspace E_λ is an invariant subspace of A, i.e., $AE_\lambda \subseteq E_\lambda$
- The dimension of E_λ can be interpreted as the maximum number of linearly independent eigenvectors that can be found, all with the same eigenvalue λ
- This number is the geometric multiplicity of λ
- Geometric multiplicity can also be described as the dimension of the null space of $A - \lambda I$ since the null space is again E_λ
- Related to the question whether a given matrix may be diagonalized by a suitable choice of coordinates

Characteristic polynomial

- The characteristic polynomial of $A \in \mathbb{C}^{m \times m}$, denoted by p_A, is the degree m polynomial

$$p_A(x) = \det(xI - A)$$

- Note p is monic (i.e., the coefficient of its degree m term is 1)

Theorem

λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$

Proof.

This follows form the definition of an eigenvalue:

$$\begin{aligned} \lambda \text{ is an eigenvalue} \quad &\iff \quad \text{there is a nonzero vector } x \text{ s.t. } \lambda x - Ax = 0 \\ &\iff \quad \lambda I - A \text{ is singular} \\ &\iff \quad \det(\lambda I - A) = 0 \end{aligned}$$

□

Characteristic polynomial (cont'd)

- Even if matrix A is real, some of its eigenvalues may be complex
- Physically, related to the phenomenon that real dynamical systems can have motions that oscillate as well as grow or decay
- Algorithmically, even if the input to a matrix eigenvalue problem is real, the output may have to be complex
- By the fundamental theorem of algebra, we can write p_A in terms of their roots

  $$p_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_m)$$

  for some numbers $\lambda_j \in \mathbb{C}$
- Algebraic multiplicity of an eigenvalue λ of A: its multiplicity as a root of p_A
- An eigenvalue is simple if is algebraic multiplicity is 1
- Algebraic multiplicity is always as great as its geometric multiplicity
- There may not be sufficient eigenvectors to span the entire space

Example

- Consider the matrices

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

- Both A and B have characteristic polynomial $(z-2)^3$, so there is a single eigenvalue $\lambda = 2$ of algebraic multiplicity 3
- For A, we can choose three independent eigenvalectors, e.g., $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and so the geometric multiplicity of $\lambda = 2$ is 3
- For B, on the other hand, we can only have one single independent eigenvector, i.e., a scalar multiple of $\mathbf{e}_1$, so the geometric multiplicity of the eigenvalue is only 1
- It means that there are not sufficient number of independent eigenvectors to span B
- It also means that A can be diagonalized but not B

Eigenvalue properties

- If $X \in \mathbb{C}^{m \times m}$ is nonsingular, then the map $A \mapsto X^{-1}AX$ is called a similarity transformation of A
- Two matrices A and B are similar if there is a similarity transformation relating one to the other, i.e., if there exists a nonsingular $X \in \mathbb{C}^{m \times m}$ s.t. $B = X^{-1}AX$
- If X is nonsingular, then A and $X^{-1}AX$ have the same characteristic polynomial eigenvalues, and algebraic and geometric multiplicities
- An eigenvalue whose algebraic multiplicity exceeds its geometric multiplicity is a defective eigenvalue
- A matrix that has one or more defective eigenvalues is a defective matrix
- Any diagonal matrix is nondefective, and both the algebraic and the geometric multiplicities of an eigenvalue λ are equal to the number of its occurrences along the diagonal

Diagonalizability

Theorem

An $m \times m$ matrix A is nondefective if and only if it has an eigenvalue decomposition $A = X\Lambda X^{-1}$

Proof.

($\Leftarrow$) Given an eigenvalue decomposition $A = X\Lambda X^{-1}$, we know Λ is similar to A (due to similarity transformation) with the same eigenvalues and the same multiplicities. Since Λ is a diagonal matrix, it is nondefective, and thus the same holds for A.

($\Rightarrow$) A nondefective matrix must have m linearly independent eigenvectors as eigenvectors with different eigenvalues must be linearly independent, and each eigenvalue can contribute as many linearly independent eigenvectors as its multiplicity. If these m independent eigenvectors re formed into the columns of a matrix X, then X is nonsingular and we have $AX = X\Lambda$, $A = X\Lambda X^{-1}$. □

Determinant and trace

Theorem

The determinant $\det(A)$ *and trace* $\mathrm{tr}(A)$ *are equal to the product and the sum of eigenvalues of* A*, respectively, counted with algebraic multiplicity*

$$\det(A) = \prod_{j=1}^{m} \lambda_j \quad \mathrm{tr}(A) = \sum_{j=1}^{m} \lambda_j$$

Proof.

$$A = U\Sigma V^T, \det(A) = \det(U)\det(\Sigma)\det(V^\top) = \prod_{j=1}^{m} \lambda_j$$

$$\begin{aligned} p_A(x) &= (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_m) \\ &= x^m - (\textstyle\sum_{j=1}^m \lambda_j)x^{m-1} + \cdots + \prod_{j=1}^m \lambda_j \end{aligned}$$

On the other hand, from characteristic polynomial

$$\begin{aligned} p_A(x) &= \det(xI - A) \\ &= x^m - (\mathrm{tr}(A))x^{m-1} + \cdots + \det(A) \end{aligned}$$

□

Hermitian matrix

- Hermitian matrix: a square matrix $A \in \mathbb{C}$ with complex entries which is equal to its own conjugate transpose

$$A_{ij} = \overline{A_{ij}}, \quad A = A^H \quad (A = A^*)$$

where A^H (or A^*) is the conjugate transpose of A, e.g.,

$$\begin{bmatrix} 3 & 2+i \\ 2-i & 1 \end{bmatrix}$$

- A complex square matrix A is normal if $A^H A = AA^H$
- A complex square matrix A is a unitary matrix if $A^H A = AA^H = I$
- A unitary matrix in which all entries are real is an orthogonal matrix
- Properties:
  - Real entries on the main diagonal
  - A matrix has only real entries is Hermitian if and only if it is a symmetric matrix
  - A real and symmetric matrix is a special case of a Hermitian matrix

Unitary and orthogonal matrices

- Unitary matrix: a complex matrix $Q \in \mathbb{C}^{m \times m}$ whose columns (or rows) constitute an orthonormal basis
  - $Q^H Q = I$
  - $Q^H Q = I \iff QQ^H = I$
  - $Q^{-1} = Q^H$
  - $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \ \forall \mathbf{x} \in \mathbb{C}^m$
- Orthogonal matrix: a real matrix $P \in \mathbb{R}^{m \times m}$ whose columns (or rows) constitute an orthonormal basis
  - $P^\top P = I$
  - $P^\top P = I \iff PP^\top = I$
  - $P^{-1} = P^\top$
  - $\|P\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \ \forall \mathbf{x} \in \mathbb{R}^m$

Complex vector and matrix

- For complex vectors, $\mathbf{x}$,
$$\|\mathbf{x}\|^2 = \mathbf{x}^H\mathbf{x}$$
- A is unitary if
$$\|A\mathbf{x}\|^2 = \mathbf{x}^H A^H A\mathbf{x} = \mathbf{x}^H\mathbf{x} = \|\mathbf{x}\|^2$$
- Two vectors are orthogonal if
$$\mathbf{x}_1^H\mathbf{x}_2 = 0$$
then $A\mathbf{x}_1$ and $A\mathbf{x}_2$ are orthogonal under unitary transformation
$$\mathbf{x}_1^H A^H A\mathbf{x}_2 = \mathbf{x}_1^H\mathbf{x}_2 = 0$$

Schur decomposition

Theorem

Every square matrix can be factorized in Schur decomposition

$$\begin{aligned} A &= QTQ^H, \quad A \in \mathbb{C}^{m\times m} \\ T &= Q^H A Q \end{aligned}$$

where Q is unitary and T is upper triangular, and the eigenvalues of A appear on the diagonal of T

- Play an important role in eigenvalue computation
- Any square matrix, defective or not, can be triangularized by unitary transformations
- The diagonal elements of a triangular matrix are its eigenvalues
- The unitary transformations preserve eigenvalues

Schur decomposition (cont'd)

Proof.

For $m = 1$, trivial case. For $m > 1$, assume that all $(m-1) \times (m-1)$ matrices are unitary similar to an upper triangular matrix, and consider an $m \times m$ matrix A. Suppose that $(\lambda, \mathbf{x})$ is an eigenpair for A and $\|\mathbf{x}\|_2 = 1$. We can construct a Householder reflector $R = R^H = R^{-1}$ with property that $R\mathbf{x} = \mathbf{e}_1$ or $\mathbf{x} = R\mathbf{e}_1$. Thus $\mathbf{x}$ is the first column in R, and so $R = \left[\mathbf{x}|V\right]$,

$$R^H AR = RA\left[\mathbf{x}|V\right] = R\left[\lambda\mathbf{x}|AV\right] = \left[\lambda\mathbf{e}_1|R^H AV\right] = \begin{bmatrix} \lambda & \mathbf{x}^H AV \\ \mathbf{0} & V^H AV \end{bmatrix}$$

Since $V^H AV$ is $(m-1) \times (m-1)$, the induction hypothesis insures that there exists a unitary matrix Q s.t. $Q^H(V^H AV)Q = \widetilde{T}$ is upper triangular. If $U = R\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q \end{bmatrix}$, then U is unitary (as $U^H = U^{-1}$), and

$$U^H AU = \begin{bmatrix} \lambda & \mathbf{x}^H AVQ \\ \mathbf{0} & Q^H V^H AVQ \end{bmatrix} = \begin{bmatrix} \lambda & \mathbf{x}^H AVQ \\ \mathbf{0} & \widetilde{T} \end{bmatrix} = T$$

is upper triangular $\square$

Diagonalization and eigenvalue problems I

- Sometimes not only $m \times m$ matrix A may have m linearly independent eigenvectors, but also they are orthogonal
- In such cases, A is unitarily diagonalizable
- A square matrix A is unitarily diagonalizable if there exists a unitary matrix Q such that

$$A = Q \Lambda Q^H$$

  where Λ is diagonal

Theorem

A Hermitian matrix is unitarily diagonalizable, and its eigenvalues are real

Theorem

A matrix is unitarily diagonalizable if and only if it is normal

Diagonalization and eigenvalue problems II

- Two matrices A and B, diagonalizable or not, are similar if they are related by

$$A = QBQ^{-1}$$

and the transformation of B into A (or vice versa) is called a similarity transformation

- If A is diagonalizable

$$\begin{aligned} AQ &= Q\Lambda \\ A\mathbf{q}_i &= \lambda_i \mathbf{q}_i \end{aligned}$$

where λ_i and $\mathbf{q}_i$ are solutions of the eigenvalue problem

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

- Derive this equation from the requirement of diagonalizing a matrix by a similarity transformation

Eigenvalue revealing factorization

- A diagonalization $A = X \Lambda X^{-1}$ exists if and only if A is nondefective
- A unitary diagonalization $A = Q \Lambda Q^H$ exits if and only if A is normal
- A unitary triangularization (Schur factorization) $A = QTQ^H$ always exists
- Will use one of these factorization to compute eigenvalues
- In general, we will use Schur factorization as this applies without restriction
- If A is normal, then Schur form comes out diagonal and its eigenvalues are real
- If A is Hermitian, then we can take advantage of symmetry with half as much work or less than is required for general A