Lecture 10 - Eigenvalues problem

Department of Computer Science
University of Houston

February 28, 2008

Introduction

- Eigenvalue problems form an important class of problems in scientific computing. The algorithms to solve them are powerful, yet far from obvious! Here we review the theory of eigenvalues and eigenvectors. Algorithms are discussed in later lectures.
- From now own, let A be square ($m \times m$).
- Let $x \neq 0 \in \boldsymbol{R}^m$.
- Then x is an eigenvector of A and $\lambda \in \boldsymbol{R}$ is its corresponding eigenvalue if $Ax = \lambda x$.
- The idea is that the action of A on a subspace S of $\boldsymbol{R}^m$ can act like scalar multiplication.
- This special subspace S is called an eigenspace.
- The set of all the eigenvalues of a matrix A is called the spectrum of A, denoted $\Lambda(A)$.

Eigenvalue decomposition

- An eigenvalue decomposition of A is a factorization

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

where X is nonsingular and I is diagonal.

- Such a decomposition does not always exist!
- The definition can be rewritten as

$$\begin{aligned} AX &= XI \\ A[x_1|x_2|\dots|x_n] &= [x_1|x_2|\dots|x_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \end{aligned}$$

- This makes it clear that

$$Ax_j = \lambda_j x_j$$

Geometric multiplicity

- The geometric multiplicity of an eigenvalue λ is the number of linearly independent eigenvectors associated with it.
- The set of eigenvectors corresponding to a single eigenvalue (plus the zero vector) forms a subspace of $\boldsymbol{R}^m$ known as an eigenspace.
- The eigenspace corresponding to $\lambda \in \Lambda(A)$ is denoted E_λ.
- E_λ is an invariant subspace of A :

$$AE_\lambda \subseteq E_\lambda$$

- The dimension of E_λ can then be interpreted as geometric multiplicity of λ.
- The maximum number of linearly independent eigenvectors that can be found for a given λ.

Characteristic polynomial

- The characteristic polynomial $p(A)$ of A is the degree-m polynomial

$$p_A(z) = det(zI - A)$$

Theorem

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

- Even in A is real, λ could be complex! However, if A is real, any complex λ must appear in complex conjugate pairs.
- If A is real and $\lambda = a + ib$ is an eigenvalue, then so is $\lambda^* = a - ib$.

Algebraic multiplicity

- The polynomial $p_A(z)$ can be written as

$$p_A(z) = (z - \lambda_1)(z - \lambda_2)\dots(z - \lambda_m)$$

  for some numbers $\lambda_j \in \mathbb{C}$ (the roots of the polynomial).
- Each λ_j is an eigenvalue of A, and in general may be repeated,

$$\lambda^2 - 2\lambda + 1 = (\lambda - 1)(\lambda - 1)$$

- The algebraic multiplicity of an eigenvalue λ as the multiplicity of λ as a root of $p_A(z)$.
- An eigenvalue is simple if its algebraic multiplicity is 1.

Theorem

If $A \in \mathbb{R}^{m\times m}$, then A has m eigenvalues counting algebraic multiplicity. In particular, if the roots of $p_A(z)$ are simple, then A has m distinct eigenvalues.

Similarity transformations (1)

- If $X \in \boldsymbol{R}^{m \times m}$ is nonsingular, then

$$A \longrightarrow X^{-1}AX$$

  is called a similarity transformation of A.
- Two matrices A and B are similar if there is a similarity transformation of one to the other.
- There is $X \in \boldsymbol{R}^{m \times m}$ is nonsingular, then

$$B = X^{-1}AX$$

- Many properties are shared by matrices that are similar.

Similarity transformations (2)

Theorem

If X is nonsingular, then A and $X^{-1}AX$ have the same characteristic polynomial, eigenvalues, and algebraic and geometric multiplicities.

Theorem

The algebraic multiplicity of an eigenvalue λ is at least as large as its geometric multiplicity.

Defective eigenvalues and matrices (1)

- A generic matrix will have algebraic and geometric multiplicities that are equal (to 1) since eigenvalues are often not repeated.
- However, this is certainly not true of every matrix!
- Consider,

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

- Both A and B have a single eigenvalue $\lambda = 2$ with algebraic multiplicity 3.

Defective eigenvalues and matrices (2)

- For A, we can choose 3 linearly independent eigenvectors, e_1, e_2, e_3. So, the geometric multiplicity of A is 3. However, for B, we only have 1 linearly independent eigenvector, e_1. So, the geometric multiplicity of B is 1.
- An eigenvalue whose algebraic multiplicity is greater than its geometric multiplicity is called a defective eigenvalue.
- A matrix that has at least one defective eigenvalue is a defective matrix i.e., it does not possess a full set of m linearly independent eigenvectors.
- Every diagonal matrix is non-defective, with algebraic multiplicity of every eigenvalue λ equal to its geometric multiplicity (equal to the number of times it occurs on the diagonal).

Diagonalizability

- Non-defective matrices are precisely those matrices that have an eigenvalue decomposition.

Theorem

$A \in \mathbb{R}^{m \times m}$ is non-defective if and only if it has an eigenvalue decomposition

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

- In view of this, another term for non-defective is diagonalizable.

Determinant and trace

- Both the trace of $A \in \boldsymbol{R}^{m\times m}$ ($tr(A) = \sum_{j=1}^{m} a_{jj}$) and its determinant are related to its eigenvalues.

Theorem

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

Orthogonal diagonalization (1)

- A may have m linearly independent eigenvectors. Sometimes hese vectors can be chosen to be orthogonal.
- In such cases we say that A is orthogonally diagonalizable; i.e., there exists an orthogonal matrix Q such that

$$A = Q\Lambda Q^T$$

- Such a decomposition is both an eigenvalue decomposition and a SVD (except possibly for the signs of the elements of λ).

Theorem

A symmetric matrix is orthogonally diagonalizable and its eigenvalues are real.

Orthogonal diagonalization (2)

- This is not the only class of orthogonally diagonalizable matrices.
- It turns out that the entire class of orthogonally diagonalizable matrices has an elegant characterization.
- We say that a matrix is normal if $AA^T = A^TA$
- . Then we have

Theorem

A matrix is orthogonally digonalizable if and only if it is normal.

Schur factorization

- This final factorization is actually the most useful in numerical analysis because all matrices (even defective ones) have a *Schur factorization*

$$A = QTQ^T$$

where Q is orthogonal and T is upper-triangular.

- Since A and T are similar, the eigenvalues of A appear on the diagonal of T.

Theorem

Every square matrix A *has a Schur factorization.*

Eigenvalue-revealing factorizations (1)

- We have just described three eigenvalue-revealing factorizations.
- Factorizations where the matrix is reduced to a form where the eigenvalues can simply be read off.
- We summarize them as follows:
  1. A diagonalization $A = X\Lambda X^{-1}$ exists if and only if A is non-defective.
  2. An orthogonal diagonalization $A = Q\Lambda Q^T$ exists if and only if A is normal.
  3. An orthogonal triangularization (Schur decomposition)$A = QTQ^T$ always exists.

Eigenvalue-revealing factorizations (2)

- To compute eigenvalues, we will construct one of these factorizations.
- In general, we will use the Schur decomposition since it applies to all matrices without restriction and it uses orthogonal transformations, which have good stability properties.
- If A is normal, then its Schur factorization will have a diagonal T.
- Moreover, if A is symmetric, we can exploit this symmetry to reduce A to diagonal form with half as much work or less than is required for general A.

Gershgorin circle

- Let A be a complex $n \times n$ matrix, with entries (a_{ij}).
- For $i \in 1, \ldots, n$ write $R_i = \sum_{j \neq i} |a_{ij}|$ where $|a_{ij}|$ denotes the complex norm of a_{ij}.
- Let $D(a_{ii}, R_i)$ be the closed disc centered at a_{ii} with radius R_i. Such a disc is called a Gershgorin disc.

Theorem

Every eigenvalue of A lies within at least one of the Gershgorin discs $D(a_{ii}, R_i)$.

Theorem

The eigenvalues of A must also lie within the Gershgorin discs C_j corresponding to the columns of A.

- For a diagonal matrix, the Gershgorin discs coincide with the spectrum. Conversely, if the Gershgorin discs coincide with the spectrum, the matrix is diagonal.