Lecture Notes 9, Math/Comp 128, Math 250

Misha Kilmer
Tufts University

October 29, 2007

More on Eigenvalues

Recall the eigenvector/eigenvalue equation:

$$Ax = \lambda x$$

- λ is a scalar (can be complex, even if A is not)
- x is the eigenvector corresponding to λ

  is never the zero vector!)

  can have complex entries (even if A is real)

Definitions

- $p(z) = \det(A - zI)$ is the characteristic polynomial
- $\det(A - zI) = 0$ is the characteristic equation, and it's roots $\lambda_1, \ldots, \lambda_m$ are the eigenvalues.
- Algebraic multiplicity of λ_i: the multiplicity of this root in the characteristic polynomial. For example, if

$$p(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m) = (z - 2)^3 (z + 1)^2 (z + 4)$$

  then 2 is an eigenvalue of algebraic multiplicity 3; -1 has algebraic multiplicity 2; the two eigenvalues $\pm 2i$ have multiplicity 1.
- Geometric multiplicity of λ_i: the dimension of the eigenspace corresponding to λ_i. Must be at least 1, at most the algebraic multiplicity.
- A defective eigenvalue is one for which the algebraic multiplicity is strictly greater than the geometric multiplicity.
- A matrix that has one or more defective eigenvalues is a defective matrix.

Important Facts

- (24.5) An $m \times m$ matrix A is nondefective iff it has an eigenvalue decomposition $A = X\Lambda X^{-1}$.
- (24.6) The determinant $det(A)$ and trace $tr(A)$ are equal to the product and sum of the eigenvalues of A, respectively, counted with algebraic multiplicity.

Unitary Diagonalization

A matrix is unitarily diagonalizable if there exists a unitary Q such that

$$A = Q\Lambda Q^*.$$

Theorem 24.7 A hermitian matrix is unitarily diagonalizable, and its eigenvalues are real.

Proof that the eigenvalues are real on board.

A matrix A is normal if $A^*A = AA^*$.

Theorem 24.8 A matrix is unitarily diagonalizable iff it is normal.

Similarity Transforms

Two matrices A and B are similar if there exists an invertible X such that $A = XBX^{-1}$.

Note that these two matrices have the same eigenvalues (with same algebraic multiplicities), but not the same eigenvectors (unless $X = I$).

We will make use of similarity transformations in the algorithms for computing eigenvalues. In particular, note that if B is upper triangular, then you know the eigenvalues of A...

Schur Factorization

A Schur Factorization of A is a factorization

$$A = QTQ^*$$

where T is upper-triangular and Q is unitary.

So A is similar to T.

Theorem 24.9 Every square matrix A has a Schur factorization! (Proof is by induction, Math 250 students should understand this)

Summary

- A diagonalization $A = X\Lambda X^{-1}$ exists iff A is nondefective.
- A unitary diagonalization $A = Q\Lambda Q^*$ exists iff A is normal.
- The Schur factorization $A = QTQ^*$ always exists.

Bottom line for designing algorithms: the last two are more ideal, as they give rise to numerically stable algorithms (because of the orthogonal Q's). IF A IS normal, then the Schur factorization comes out diagonal anyway. IF A IS hermitian, can take advantage and reduce amt. of work to 1/2.

One exercise

True or False: If λ is an eigenvalue of A, and c is a real number, then $\lambda - c$ is an eigenvalue of $A - cI$.

True: (pf. on board.)

Note that the eigenvectors are unchanged!

Hermitian Positive Definite Matrices

Let A be real and symmetric positive definite. Then all the eigenvalues are real AND positive!

Likewise, if A is complex, hermitian and positive definite.

(Proof positive, on board.)

Reminder: Singular Matrices

Recall that an $m \times m$ matrix A is singular (non-invertible) iff 0 is an eigenvalue of A.

WHY?

Beginning Computations

The “obvious” algorithm: Find the roots of the characteristic polynomial.

Cursed by degree, and by ill-posedness of the rootfinding problem.

Theorem 25.1 For any $m \geq 5$, there is a polynomial of degree m with rational coefficients that has a real root with the property that it cannot be written using any expression involving rational numbers, addition, subtraction, division, and kth roots.

Consequences

- Even in exact arithmetic, we couldn’t produce the exact roots of an arbitrary polynomial in a finite number of steps.
- Any eigenvalue solver must be iterative (and avoid the root finding problem, when necessary)

Next Obvious Algorithm

The Power Iteration.

Let's assume that $|\lambda_1| > |\lambda_2| > |\lambda_3| > \cdots |\lambda_m| > 0$, so that there is a single, largest magnitude eigenvalue, and all the eigenvalues are distinct. Then there is a complete set of eigenvectors, $x_1, \ldots, x_m$, and this set spans all of $\mathbb{C}^m$. We can assume the eigenvectors have unit length.

So let v be any vector (assume it has unit length) in $\mathbb{C}^m$. Expand $v = \sum_{i=1}^m c_i x_i$. (We don't know the c_i's, but it doesn't matter as long as we know $c_1 \neq 0$.)

Set

$$v^{(1)} = Av = \sum_{i=1}^m c_i A x_i = \sum_{i=1}^m c_i \lambda_i x_i,$$

$$v^{(2)} = Av^{(1)} = A^2 v = \sum_{i=1}^{m} c_i \lambda_i A x_i = \sum_{i=1}^{m} c_i \lambda_i^2 x_i$$

$$v^{(m)} = A^m v = \sum_{i=1}^{m} c_i \lambda_i^m x_i$$

Clearly, the dominant term is the first in the sum, for large enough m.

If we normalize at each step $w^{(m)} \leftarrow \frac{v^{(m)}}{\|v^{(m)}\|}$ (see symmetric A, for example), the sequence converges to the eigenvector corresponding to the largest eigenvalue of A.

General, Two-Step Process

Preliminaries: A matrix is called upper Hessenberg if the everything below the first subdiagonal is zero.

Example: A 5×5 upper Hessenberg matrix

$$\begin{bmatrix} 4 & 1.5 & 2 & 0 & 3.7 \\ 1 & 3.2 & 5 & 7 & 1.4 \\ 0 & 9 & 11 & 3 & 6 \\ 0 & 0 & 3 & 2.3 & 8 \\ 0 & 0 & 0 & 6.2 & 1 \end{bmatrix}$$

Note that a symmetric (Hermitian) $n \times n$ upper Hessenberg must be tridiagonal.

Two-Step Process

For a *non-symmetric* (non-Hermitian) matrix

- Phase-I Reduce the matrix to upper Hessenberg form with orthogonal similarity transforms

  Householder reflectors

  Givens rotations

  Requires $O(m^3)$ flops, is a "direct" process

- Phase-2 Reduce the upper Hessenberg to upper triangular form with similarity transforms

  this is iterative, never runs to completition, but converges to the machine precision in $O(m)$ iterations. Each iteration costs $O(m^2)$ flops, therefore, $O(m^3)$ for this phase applied to the upper Hessenberg matrix.

If Phase-2 were applied directly to a non-reduced form, it would cost $O(m^4)$ flops!!

Two-Step Process

For a symmetric (Hermitian A),

- Phase-I Reduce the matrix to tridiagonal form with orthogonal similarity transforms

  Householder reflectors

  Givens rotations

  Requires $O(m^3)$ flops, is a "direct" process

- Phase-2 Reduce the upper Hessenberg to diagonal form with similarity transforms

  this is iterative, never runs to completitition, but converges to the machine precision in $O(m)$ flops.

Phase I

Why can't we reduce to upper triangular form directly with similarity transforms??

See board.

Phase I

Why upper Hessenberg (tridiagonal) form??

See board.