Numerical Analysis

Daming Li

Department of Mathematics, Shanghai JiaoTong University,
Shanghai, 200240, China
Email: lidaming@sjtu.edu.cn

October 14, 2014

Let A be an $n \times n$ matrix (whose elements may be complex number). Let λ be a a scalar (complex number). If

$$Ax = \lambda x$$

has a nontrivial solution $((x \neq 0))$, then λ is an eigenvalue of A. Consider

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 5 & -1 & 2 \\ -3 & 2 & -5/4 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ -4 \end{pmatrix} = -2 \begin{pmatrix} 1 \\ 3 \\ -4 \end{pmatrix}$$

The eigenvalues satisfies the characteristic equation of the matrix A

$$\det(A - \lambda I) = 0$$

Power Method

To illustrate these concepts, the characteristic equation of

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 2 & 1 & 1 \end{pmatrix} \quad \text{is}$$

$$\begin{aligned} & \det \begin{pmatrix} 1-\lambda & 2 & 1 \\ 0 & 1-\lambda & 3 \\ 2 & 1 & 1-\lambda \end{pmatrix} \\ = \quad & -\lambda^3 + 3\lambda^2 + 2\lambda + 8 = -(\lambda - 4)(\lambda^2 + \lambda + 2) \end{aligned}$$

The eigenvalues are

$$4, \quad -\frac{1}{2} \pm \frac{\sqrt{7}}{2} i$$

For large system, the solving of characteristic equation is not recommended. This is because the roots of the characteristic equation is rather sensitive to the coefficients of characteristic equation.

Power Method

The power method is designed to compute the dominant eigenvalue and its corresponding eigenvector. Assume that A has two properties: (1) there is a single eigenvalue of maximum modulus, (2) there is n independent set of eigenvectors. The eigenvalues can be labeled so that

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$$

and the corresponding eigenvectors $u^{(1)}, u^{(2)}, \ldots, u^{(n)}$

$$Au^{(i)} = \lambda_i u^{(i)} \quad (1 \leq i \leq n)$$

Let $x^{(0)} \in C^n$, which can be written as

$$x^{(0)} = a_1 u^{(1)} + \cdots a_n u^{(n)} \quad (a_1 \neq 0)$$

Power Method

We form

$$x^{(k)} = Ax^{(k-1)}, \quad k = 0, 1, \cdots$$

so that

$$x^{(k)} = A^k x^{(0)} = \sum_{i=1}^{n} a_i \lambda_i^k u^{(k)} = \lambda_1^k [a_1 u^{(1)} + \varepsilon^{(k)}]$$

where

$$\varepsilon^{(k)} = \sum_{i=2}^{n} a_i (\lambda_i/\lambda_1)^k x_i \to 0, \quad k \to \infty$$

Let ϕ be any linear function in C^n which satisfies $\phi(u^{(1)}) \neq 0$. Then

$$\phi(x^{(k)}) = \lambda_1^k [a_1 \phi(u^{(1)}) + \phi(\varepsilon^{(k)})]$$

Consequently, the following ratios converge to λ_1

$$r_k \equiv \frac{\phi(x^{(k+1)})}{\phi(x^{(k)})} = \lambda_1 \frac{a_1 \phi(u^{(1)}) + \phi(\varepsilon^{(k+1)})}{a_1 \phi(u^{(1)}) + \phi(\varepsilon^{(k)})} \to \lambda_1$$

Power Method

This constitutes the power method for computing λ_1. Since the direction of $x^{(k)}$ aligns more and more with $u^{(1)}$ as $k \to \infty$, the method can also give us the eigenvector $u^{(1)}$.

If λ is an eigenvalue of A, λ^{-1} is an eigenvalue of A^{-1}. Thus the power method applying A^{-1} can be used to compute the smallest eigenvalue of A. Suppose that

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

This implies that A is nonsingular. The eigenvalues of A^{-1} are arranged to be

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

Consequently, we can compute λ_1^{-1} by applying the power method to A^{-1}. It is not a good idea to calculate A^{-1}. Rather, we obtained $x^{(k+1)}$ by

$$Ax^{(k+1)} = x^{(k)}$$

This method is called the inverse power method.

We have outline the power method for computing the largest eigenvalue and the inverse power method for computing the smallest eigenvalue. Suppose that these exists a complex number μ such that

$$0 < |\lambda_k - \mu| < \varepsilon, \quad |\lambda_j - \mu| \geq \varepsilon, \quad j \neq k$$

Since the eigenvalues of $A - \mu I$ is $\lambda_i - \mu$, the inverse power method can be applied to $A - \mu I$, resulting in an approximate of $(\lambda_k - \mu)^{-1}$

$$(A - \mu I)x^{(k+1)} = x^{(k)}$$

This algorithm is called the shifted inverse power method. Thus this method is used to calculate the eigenvalue closest to μ.

Similarly, We can calculate the eigenvalue, say λ_k, farthest from a given value μ. Suppose that

$$|\lambda_k - \mu| > \varepsilon, \quad |\lambda_j - \mu| < \varepsilon, \quad j \neq k$$

The power method applied to $(A - \mu I)$ computes $\lambda_k - \mu$, and thus the approximate of λ_k can be obtained. This algorithm is called the shifted power method.

Exercises: P.234

1,2,3,8,16

The two matrices A and B are similar to each other, if there exists a nonsingular matrix P such that $B = PAP^{-1}$.
Similar matrices have the same eigenvalues.
Proof. Let A and B are simular matrices, that is

$$B = PAP^{-1}$$

then

$$\det(B - \lambda I) = \det(PAP^{-1} - \lambda I) = \det(A - \lambda I)$$

A matrix U is called unitary if $UU^* = I$.
Every square matrix is unitarily similar to a trianglar matrix.

Proof. We proceed by induction on n, the order of the matrix A. The theorem is trivial for $n = 1$. Suppose that the theorem has been proved for $n - 1$ and consider the matrix of order n. Let λ be eigenvalue and x be the corresponding eigenvector with $\|x\|_2 = 1$. We have a unitary matrix U such that $Ux = \beta e^{(1)} \equiv \beta(1, 0, \cdots, 0)^T$ where $\beta = x_1/|x_1|$ if $x_1 \neq 0$ and $\beta = 1$ if $x_1 = 0$. Thus

$$UAU^* e^{(1)} = UA\beta^{-1}x = \beta^{-1}\lambda Ux = \lambda e^{(1)}$$

This proves that the first column of UAU^* is $\lambda e^{(1)}$. Let $\tilde{A}$ denote the matrix obtained by the first row and column of UAU^*. By the induction hypothesis, there is a unitary matrix Q of order $n - 1$ such that $Q\tilde{A}Q^*$ is triangular.

The unitary matrix that reduces A to triangular form is

$$V = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} U$$

because

$$\begin{aligned} VAV^* &= \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} UAU^* \begin{pmatrix} 1 & 0 \\ 0 & Q^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} \lambda & w \\ 0 & \tilde{A} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q^* \end{pmatrix} \\ &= \begin{pmatrix} \lambda & wQ^* \\ 0 & Q\tilde{A}Q^* \end{pmatrix} \end{aligned}$$

Since the matrix is unitary, the prof is complete.

Schur's and Gershgorin's Theorem

Every square matrix is similar to a triangular matrix.
Every Hermite matrix is unitarily similar to a diagonal matrix.
Proof. If A is Hermitian, then $A = A^*$. Let U be a unitary matrix such that UAU^* is upper triangular. Then $(UAU^*)^*$ is lower triangular. But

$$(UAU^*)^* = UAU^*$$

Thus the matrix UAU^* is both upper and lower triangular; hence, it must be a diagonal matrix.
The matrix $I - vv^*$ is unitary if and only if $\|v\|_2^2 = 2$ or $v = 0$.
Proof.

$$I = UU^* = (I - vv^*)(I - vv^*) = I - 2vv^* + (v^*v)vv^* = I - (2 - v^*v)vv^*$$

Let x and y be two vectors such that $\|x\|_2 = \|y\|_2$ and (x, y) is real. Then there exists a unitary matrix U of the form $I - vv^*$ such that $Ux = y$.

Proof. If $x = y$, let $v = 0$. If $x \neq y$, let $v = \alpha(x - y)$, with $\alpha = \sqrt{2}/\|x - y\|_2$.

$$\begin{aligned} Ux - y &= (I - vv^*)x - y = x - y - \alpha^2(x - y)(x^* - y^*)x \\ &= (x - y)[1 - \alpha^2(x^*x - y^*x)] = 0 \end{aligned}$$

by noting the hypotheses $x^*x = y^*y$ and $y^*x = x^*y$ and

$$2(x^*x - y^*x) = x^*x + y^*y - y^*x - x^*y = (x^* - y^*)(x - y) = \|x - y\|_2^2$$

If the eigenvalue of λ of an $n \times n$ matrix A is known, then the proof of Schur's theorem shows how to produce an $(n-1) \times (n-1)$ matrix $\tilde{A}$ whose eigenvalues are the same as those of A, except for λ. This procedure is known as deflation.
The formalization of the deflation for eigenvalues is as follows:

- Obtain an eigenvector x corresponding to a known eigenvalue A.
- Define $\beta = x_1/|x_1|$, if $x_1 \neq 0$, and let $\beta = 1$ otherwise
- Define $\alpha = \sqrt{2}/\|x - \beta e^{(1)}\|_2$, $v = \alpha(x - \beta e^{(1)})$, and $U = I - vv^*$
- Let $\tilde{A}$ be the matrix obtained by the first row and column of UAU^*

Schur's and Gershgorin's Theorem

The spectrum of an $n \times n$ matrix A (that is, the set of its eigenvalues) is contained in the union of the following n disks, D_i, in the complex plane:

$$D_i = \left\{x \in C : |z - a_{ii}| \leq \sum_{i \neq j = 1}^{n} |a_{ij}|\right\} \quad (1 \leq i \leq n)$$

Proof. $Ax = \lambda x$ with $||x||_\infty = 1$, Let i be an index for which $|x_i| = 1$. We have

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j$$

Therefore

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

and

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

Exercises: P.242

1,2,4,12,14,29

Householder's QR Factorization

We turn now to the most useful of orthogonal factorizations, namely, one due to Alston Householder and known by his name. The objective is to factor a$m \times n$ matrix A into a product

$$A = QR$$

where Q is an $m \times m$ unitary matrix and R is an $m \times n$ upper triangular matrix. The factorization algorithm actually produces

$$Q^* A = R$$

and Q^* is built up step by step as a product of unitary matrices having the special form

$$\begin{pmatrix} I_k & 0 \\ 0 & I_{m-k} - vv^* \end{pmatrix}$$

These are called reflections or Householder transformations.

Let the original first column of A be denoted by A_1. We want $(I - vv^*)A_1 = \beta e^{(1)}$. First, select a complex number β such that $|\beta| = \|A_1\|_2$ and such that $(A_1, \beta e^{(1)})$ is real. The put $v = \alpha(A_1 - \beta e^{(1)})$ with $\alpha = \sqrt{2}/\|A_1 - \beta e^{(1)}\|_2$. This description admits two choices for β, and we select the one for which there is less cancellation in computing the first component of v. To understand how this is done, write

$$\beta = \|A_1\|_2 e^{i\phi}, \quad a_{11} = |a_{11}| e^{i\theta}$$

Then we have

$$(A_1, \beta e^{(1)}) = a_{11}\bar{\beta} = |a_{11}|\,\|A_1\|_2 e^{i(\theta-\phi)}$$

This must be real, and so $\theta - \phi$ should be either 0 or π.

If we choose $\theta - \phi = \pi$, the first component of v will have no subtractive cancellation, since

$$v_1 = \alpha(a_{11} - \beta) = \alpha(|a_{11}| + |\beta|)e^{i\theta}$$

Thus, we define β by

$$\beta = -\|A_1\|_2 e^{i\theta} = -\|A_1\|_2 a_{11}/|a_{11}|$$

The algorithm for obtaining the matrix U in this first step is as follows.

$$\beta \leftarrow -\|A_1\|_2 a_{11}/|a_{11}|$$

$$y \leftarrow A_1 - \beta e^{(1)}, \quad \alpha \leftarrow \sqrt{2}/\|y\|_2$$

$$v \leftarrow \alpha y, \quad U \leftarrow I - vv^*$$

Householder's QR Factorization

The succeeding steps in the QR factorization are similar to the first step. After k steps, we shall have multiplied A on the left by k unitary matrices

$$U_k U_{k-1} \cdots U_1 A = \begin{pmatrix} J & H \\ 0 & W \end{pmatrix}$$

in which J is an upper triangular $k \times k$ matrix.
Construct $v \in C^{m-k}$ such that $I - vv^*$ is unitary matrix of order $m - k$ and such that $(I - vv^*)W$ has zeros below the initial element in its first column.

$$\begin{pmatrix} I & 0 \\ 0 & I - vv^* \end{pmatrix} \begin{pmatrix} J & H \\ 0 & W \end{pmatrix} = \begin{pmatrix} J & H \\ 0 & (I - vv^*)W \end{pmatrix}$$

The first factor on the left is unitary, and denoted by U_{k+1}.

After $n-1$ steps,

$$Q^*A = R \Longleftrightarrow A = QR$$

where R is $m \times n$ upper triangular matrix and

$$Q^* = U_{n-1} \cdots U_1$$

which leads

$$Q = U_1^* \cdots U_{n-1}^* = U_1 \cdots U_{n-1}$$

Here

$$U_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & I_{n-k+1} - vv^* \end{pmatrix}$$

is symmetric.

Householder's QR Factorization

An important application of the orthogonal factorizations being discussed here is to the least-squares problem for a linear system of equations. Consider a system of m equations in n unknowns written as $Ax = b$. Assume that $m \geq n = rank(A)$.
In such cases it is often required to find an x that minimizes the norm of the residual vector, $b - Ax$. The least-squares solution of $Ax = b$ is the vector x that makes $||b - Ax||_2$ a minimum.
Under the hypothesis we have made concerning the rank of A, this x will be unique.

Householder's QR Factorization

If x is a point such that $A^*(Ax - b) = 0$, then x solves the least-squares problem.

Proof. Let y be any other point. Since $A^*(Ax - b) = 0$, we conclude that $b - Ax$ is orthogonal to the column space of A. Moreover, since $A(x - y)$ is in the column space of A, we have $(b - Ax, A(x - y)) = 0$, and the Pythagorean rule gives

$$\|b-Ay\|_2^2 = \|b-Ax+A(x-y)\|_2^2 = \|b-Ax\|_2^2+\|A(x-y)\|_2^2 \geq \|b-Ax\|_2^2$$

The equation $A^*(Ax - b) = 0$ is called normal equations.

Exercises: P.255

1,9,16,17,19, 33

The Schur's Theorem shows that any matrix is unitarily similar to a triangular matrix, whose eigenvalues can be read off from the diagonal. The problem is that finding this factorization is as difficult as finding all the complex roots of a characteristic polynomial. We factorization $A = QR$ where Q is unitary and R is upper triangular. We wish R to have a nonnegative diagonal. This is easily arranged. In fact, $A = (QD)(D^*R) = \hat{Q}\hat{R}$ The definition of $D = diag(d_{ii})$ should be $d_{ii} = r_{ii}/|r_{ii}$ if $r_{ii} \neq 0$, $d_{ii} = 1$, otherwise.

The QR-algorithm of Francis is an iterative procedure

$$A_k = Q_k R_k, \quad A_{k+1} = R_k Q_k, \quad k = 1, 2, \cdots, \quad A_1 = A$$

where Q_k is unitary and R_k is upper triangular with nonnegative diagonal. Obviously, A_k is similar to A_{k+1}

$$A_k = Q_k R_k = (Q_k R_k)(Q_k Q_k^*) = Q_k A_{k+1} Q_k^*$$

Note that if the matrix A is real, then the subsequent matrix A_k will also be real. Thus, if A has some nonreal eigenvalues, we can only expect, under the best of circumstances, that A_k will converge to a "triangular" matrix with 2×2 submatrices on its diagonal.

Householder's QR Factorization

To economize on the amount of arithmetic involved in the QR-iteration, the matrix A is first reduced to upper Hessenberg form by means of unitary similarity transforms. An upper Hessenberg matrix is a matrix H in which $h_{ij} = 0$ when $i > j + 1$. The reduction of A to H by unitary similarity transforms uses an algorithm of Householder. Let us describe the kth step. At the beginning of step k, columns 1 to $k - 1$ will have the correct form for an upper Hessenberg matrix. Let us partition the partially reduced matrix in the following way,

$$\begin{pmatrix} B & C \\ D & E \end{pmatrix}$$

where B is a $k \times k$ upper Hessenberg matrix. The matrix D is $(n - k) \times k$, and has zeros everywhere but in its kth column.

Let U be any unitary matrix of order $n - k$.

$$\begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}\begin{pmatrix} B & C \\ D & E \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & U^* \end{pmatrix} = \begin{pmatrix} B & CU^* \\ UD & UEU^* \end{pmatrix}$$

We shall select U so that UD will have in its kth column a vector $(\beta, 0, 0, ..., 0)^T$.
Notice that the form of the matrix D is of order $n - k$.

$$\begin{pmatrix} 0 & \cdots & 0 & d_1 \\ 0 & \cdots & 0 & d_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & d_{n-k} \end{pmatrix}$$

Householder's QR Factorization

It therefore suffices to determine U so that

$$Ud = \beta e^{(1)}, \quad d = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-k} \end{pmatrix}$$

We see that β should be chosen so that $(d, \beta e^{(1)})$ is real and so that $\|\beta e^{(1)}\|_2 = \|d\|_2$. We put $U = I - vv^*$ where $v = \alpha(d - \beta e^{(1)})$ having chosen $\beta = -(d_1/|d_1|)\|d\|_2$ and $\alpha = \sqrt{2}/\|d - \beta e^{(1)}\|_2$.

Householder's QR Factorization

The next technique that is combined with the basic QR-algorithm is a repeated origin shift. Before we discuss that technique, we give an indication of why it is necessary.
The slow convergence of the basic algorithm is alleviated by shifts performed on the successive matrices, a shift being defined as replacement of a matrix A by $A - zI$. The shifted QR-algorithm proceeds in the following manner.

$$A_1 = Hessenberg(A)$$

$$A_k - z_k I = Q_k R_k, \quad A_{k+1} = R_k Q_k + z_k I, \quad k = 1, 2, \cdots, \quad A_1 = A$$

First, A_{k+1} is upper Hessenberg if A_{k+1} is upper Hessenberg matrix. Secondly, $A_{k+1} = Q_k^* A_k Q_k$ is similar to A_k. Thirdly, if the scalar z_k in the algorithm is taken to be the lower right diagonal element of A_k, then the iteration should rapidly produce a vector of the form $(0, 0, ..., 0, \alpha)^T$ in the last row, where α is then an eigenvalue of A. The best way to proceed thereafter is to deflate the matrix by dropping the last row and column in A_k.

Let A be a matrix in partitioned form

$$\begin{pmatrix} B & C \\ 0 & E \end{pmatrix}$$

in which B and E are square matrices. Then the spectrum of A is the union of the spectra of B and E.

Exercises: P.276

1,2,6,7,11