MA ASHISHA SIR

05 April 2024 02:16 AM

$Ax = b$,

Theorem 4.1. *If A is real matrix of order $n \times n$, then the following statements are equivalent.*

- *$Ax = 0$ has only trivial solution.*
- *For each b, $Ax = b$ has a solution.*
- *A is invertible.*
- $\det(A) \neq 0$.

$A^{-1} = \mathrm{Adj}(A)/\det(A)$.

4.3. Doolittle's Method

In this method A is decomposed as $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}, \qquad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}.$$

So that

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = LU = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ l_{21}u_{11} & l_{21}u_{12} + u_{22} & l_{21}u_{13} + u_{23} \\ l_{31}u_{11} & l_{31}u_{12} + l_{32}u_{22} & l_{31}u_{13} + l_{32}u_{23} + u_{33} \end{bmatrix} \tag{4.7}$$

Now one has to solve nine equations to find the all the total nine unknown coefficients of lower triangular matrix L and upper triangular matrix U. But these are easy to solve more or less only substitutions are needed. The solution is obtained by first solving $Lz = b$ for z by direct methods and then solving $Ux = z$ for x again by direct methods.

4.4. Crout's Method

Here one decomposes $A = LU$, where

$$L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}, \qquad U = \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus one has

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = LU = \begin{bmatrix} l_{11} & l_{11}u_{12} & l_{11}u_{13} \\ l_{21} & l_{21}u_{12} + l_{22} & l_{21}u_{13} + l_{22}u_{23} \\ l_{31} & l_{31}u_{12} + l_{32} & l_{31}u_{13} + l_{32}u_{23} + l_{33} \end{bmatrix} \tag{4.8}$$

Here we again need to solve nine equations to determine all the nine unknown coefficients. And similar to previous method we first solve $Lz = b$ for z and then $Ux = z$ for x.

A real square matrix A is said to be positive definite if $\det A > 0$ and all leading principal minors are positive.

4.6. A matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

is positive definite if

- $a_{11} > 0$,
- $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} > 0$.
- $\det A > 0$.

Cholesky's Method

Cholesky's method is applicable for symmetric and positive definite matrix A. In this case the decomposition of A is $A = LL^T$, where

$$L = \begin{bmatrix} d_1 & 0 & 0 \\ l_{21} & d_2 & 0 \\ l_{31} & l_{32} & d_3 \end{bmatrix}.$$

How to decompose square?

So that

$$A = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = LU = \begin{bmatrix} d_1^2 & d_1 l_{12} & d_1 l_{31} \\ d_1 l_{21} & l_{21}^2 + d_2^2 & l_{21}l_{31} + d_2 l_{32} \\ d_1 l_{31} & l_{31}l_{21} + l_{32}d_2 & l_{31}^2 + l_{32}^2 + d_3^2 \end{bmatrix} \tag{4.9}$$

Here we only need to solve six equations in six unknowns. To solve $Ax = b$ we first solve $Lz = b$, for z and then $L^T x = z$ for x.

Suppose i_{th} row is multiplied by a non-zero constant c. We denote this transformation by $R_i(c)$. If we multiply the i_{th} row of the identity matrix by the same constant c and name this new matrix by $E_{i(c)}$, then it is easy to observe that the inverse of this matrix is $E_{i(c^{-1})}$.

The row transformation of interchanging the i_{th} row with the j_{th} row is denoted by $R_i \leftrightarrow R_j$. If we interchange the i_{th} row with the j_{th} row of the identity matrix and obtain a new matrix denoted by $E_{i \leftrightarrow j}$. This is a self inverse matrix. And the same row transformation can also

Suppose we multiply a constant c to i_{th} row and add it to j_{th} row. This row transformation is denoted by $R_{ji(c)}$. The corresponding elementary matrix obtained from identity by the same row transformation is denoted by $E_{ji(c)}$. The inverse of this elementary matrix is $E_{ji(-c)}$.

Echloen form.

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$$

Row Echleon Form

$$\begin{bmatrix} 1 & a & b & c \\ 0 & 1 & d & e \\ 0 & 0 & 1 & f \end{bmatrix}$$

Row Reduced Echleon form.

$$\begin{bmatrix} 1 & 0 & b & c \\ 0 & 1 & d & e \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

pivot columns only 1 entry

Gauss Elimination

$Ax = b$

$[A \mid b] \xrightarrow{\text{Convert}} [E \mid b_T]$

echloen form

Gauss Jordan Elimination

$Ax = b$

$[A \mid b] \xrightarrow{\text{Convert}} [I \mid E]$ $E = A^{-1}$

row reduced echloen form.

Partial Pivoting

Largest modulus value in a column gets shifted upwards by row transformations.

Total Pivoting

Largest number in the whole matrix gets shifted by column transformations

NORM

$\|x\| = 0$ iff $x = 0$ else $\|x\| > 0$

$\|\alpha x\| = |\alpha| \, \|x\|$

$\|u + v\| \leq \|u\| + \|v\|$

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}. \qquad \|x\|_\infty = \max_{1 \leq i \leq \infty} \{|x_i|\}.$$

Matrix Norm

Matrix Norm

- $\|A\| \geq 0$ and $\|A\| = 0$ if and only if A is null matrix, for all $A \in M_{mn}$.
- $\|\alpha A\| = |\alpha|\|A\|$, for any scalar $\alpha \in \mathbb{R}$ and $A \in M_{mn}$.
- **Triangle Inequality:** For any $A, B \in M_{mn}$,

$$\|A + B\| \leq \|A\| + \|B\|. \tag{4.16}$$

- This norm is also compatible with multiplication by a column vector of order $n \times 1$

$$\|Ax\| \leq \|A\|\|x\|, \tag{4.17}$$

$$\|A\|_1 = \max_{1 \leq j \leq m} \left\{ \left(\sum_{i=1}^{n} |a_{ij}| \right) : j = 1, \ldots, m \right\}.$$

$$\|A\|_2 = \left(\sum_{i,j=1}^{m,n} |a_{ij}|^2 \right)^{1/2}.$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \left\{ \left(\sum_{j=1}^{m} |a_{ij}| \right) : i = 1, \ldots, n \right\}.$$

Gauss Jacobi

convert $A = L + U + D$

$$Dx = (L+U)x + b \rightarrow eq^n$$

$$x = \underbrace{D^{-1}(L+U)}_{B} x + \underbrace{D^{-1}b}_{C}$$

$$x^{K+1} = Bx^K + C$$

$$e_i^{(k+1)} = \sum_{j=1, j \neq i}^{n} \frac{-a_{ij}}{a_{ii}} e_j^{(k)}$$

$$\alpha_i = \sum_{j=1}^{i-1} \frac{|a_{ij}|}{|a_{ii}|}, \qquad \beta_i = \sum_{j=i+1}^{n} \frac{|a_{ij}|}{|a_{ii}|},$$

$$\mu = \max_{1 \leq i \leq n} \{(\alpha_i + \beta_i) : i = 1, \ldots, n\}.$$

$$\begin{aligned} |e_i^{(k+1)}| &\leq \|e^{(k)}\|_\infty \sum_{j=1, j \neq i}^{n} \frac{|a_{ij}|}{|a_{ii}|} \\ &\leq \|e^{(k)}\|_\infty \, (\alpha_i + \beta_i) \\ &\leq \|e^{(k)}\|_\infty \, \mu. \end{aligned}$$

Above inequality is true for all $i = 1, \ldots, n$, hence

$$\begin{aligned} \max_{1 \leq i \leq n} \{|e_i^{(k+1)}| : i = 1, \ldots, n\} &\leq \|e^{(k)}\|_\infty \, \mu, \\ \|e^{(k+1)}\|_\infty &\leq \mu \, \|e^{(k)}\|_\infty, \end{aligned} \tag{4.35}$$

Above inequality (4.35) is true for all $k \in \mathbb{N}$. And hence by repeated application of (4.35), we have

$$\|e^{(k+1)}\|_\infty \leq \mu\|e^{(k)}\|_\infty \leq \mu\mu\|e^{(k-1)}\|_\infty \leq (\mu)^{k+1}\|e^{(0)}\|_\infty, \tag{4.36}$$

$$\begin{aligned} & (\alpha_i + \beta_i) < 1 \quad \text{for all} \quad i = 1, \ldots, n \\ \text{or} \quad & \left(\sum_{j=1}^{i-1} \frac{|a_{ij}|}{|a_{ii}|} + \sum_{j=i+1}^{n} \frac{|a_{ij}|}{|a_{ii}|} \right) < 1 \quad \text{for all} \quad i = 1, \ldots, n \\ \text{or} \quad & \left(\sum_{j=1, j \neq i}^{n} |a_{ij}| \right) < |a_{ii}| \quad \text{for all} \quad i = 1, \ldots, n \end{aligned} \tag{4.38}$$

cond^n for solution.

The condition (4.38) is known as strict row diagonally dominant and also implies the first condition. Thus for the convergence of Gauss-Jacobi iteration method we only need the coefficient matrix to be strict row diagonally dominant.

in this we take initial guess & update after each iteration.

Remark 4.4. Note that $\|B\|_\infty = \mu$. Then from (4.31)

$$\|e^{(k)}\| = \|x - x^{(k)}\| \leq \frac{\mu^k}{1-\mu}\|x^{(1)} - x^{(0)}\|.$$

any strictly row diagonally dominant matrix is non singular

Gauss - Seidal Approach.

$$A = L + D + U$$

$$(L+D)\,x = -U\,x + b$$

$$x^{k+1} = -(L+D)^{-1} U\, x^k + (L+D)^{-1} b$$

$$e_i^{(k+1)} = \sum_{j=1}^{i-1} \frac{-a_{ij}}{a_{ii}} (e_j^{(k+1)}) + \sum_{j=i+1}^{n} \frac{-a_{ij}}{a_{ii}} (e_j^{(k)}),$$

$$\eta = \max_{1 \le i \le n} \left\{ \frac{\beta_i}{(1-\alpha_i)} : i = 1, \ldots, n \right\}.$$

$$\|e^{(k+1)}\|_\infty \le \eta \|e^{(k)}\|_\infty \le \eta\eta \|e^{(k-1)}\|_\infty \le (\eta)^{k+1} \|e^{(0)}\|_\infty.$$

$$\eta = \max_{1 \le i \le n} \left\{ \frac{\beta_i}{(1-\alpha_i)} : i = 1, \ldots, n \right\} < 1. \tag{4.48}$$

Let us first consider the third condition (4.48), which will be valid if and only if

$$\begin{aligned} & \left(\frac{\beta_i}{(1-\alpha_i)} \right) < 1 \quad \text{for all} \quad i = 1, \ldots, n \\ \text{or} \quad & \beta_i < (1-\alpha_i) \quad \text{for all} \quad i = 1, \ldots, n \\ \text{or} \quad & \left(\sum_{j=1}^{i-1} \frac{|a_{ij}|}{|a_{ii}|} + \sum_{j=i+1}^{n} \frac{|a_{ij}|}{|a_{ii}|} \right) < 1 \quad \text{for all} \quad i = 1, \ldots, n \\ \text{or} \quad & \left(\sum_{j=1, j \ne i}^{n} |a_{ij}| \right) < |a_{ii}| \quad \text{for all} \quad i = 1, \ldots, n \end{aligned} \tag{4.49}$$

Thus if we assume the coefficient matrix to be strict row diagonally dominant, the third assumption will be satisfied. Further in this case $1 - \alpha_i > \beta_i \ge 0$ implies the second condition and first is obviously true.

in this we update straight after getting the value.

Exercise 4.3. If $\mu < 1$, prove that $\eta \le \mu$.

Theorem 4.2. *Suppose A and B are two square matrices of order $n \times n$. If A is invertible and*

$$\|A - B\| < \frac{1}{\|A^{-1}\|},$$

then B is also invertible.

$$\frac{\|x\|}{\|A^{-1}\|} = \frac{\|A^{-1}Ax\|}{\|A^{-1}\|} \le \|Ax\| = \|Ax - Bx\| \le \|A - B\| \times \|x\|.$$

$$\frac{\|\delta x\|}{\|x\|} \le \|A^{-1}\| \times \|\delta b\| \frac{1}{\|x\|} \le \|A^{-1}\| \times \|\delta b\| \frac{\|A\|}{\|b\|} = (\|A^{-1}\| \times \|A\|) \frac{\|\delta b\|}{\|b\|}. \tag{4.53}$$

Thus we see that the relative error in x is controlled by relative error in b if one has control over the quantity $(\|A^{-1}\|\|A\|)$, which is known as condition number of matrix A.

Exercise 4.4. Show that for any invertible matrix A the condition number is always greater or equal to 1, when the considered norm is infinity norm.

Problem 4.1. Show that a strictly row diagonally dominant matrix is invertible.

Solution. Let A be a strictly row diagonally dominant matrix. If $A = L+D+U$ is the decomposition of A, then the diagonal matrix D consists of non zero entries in the diagonal and hence invertible. Since $A = D \times D^{-1}A$, it is sufficient to prove that $D^{-1}A$ is invertible. Note that I is invertible matrix with $\|I^{-1}\|_\infty = 1$. Thus if we can show that $\|I - D^{-1}A\|_\infty < 1$, then by the application of the previous theorem it follows that $D^{-1}A$ is invertible. But it is easy to see that $\|I - D^{-1}A\|_\infty = \mu < 1$.

EIGENVALUES

$$By = \lambda y.$$

This is equivalent to say that $(B - \lambda I)y = 0$ for some non-zero vector y,
$\Leftrightarrow$ the null space of $(B - \lambda I)$ is not equal to $\{0\}$,
$\Leftrightarrow$ the dimension of the null space of $(B - \lambda I)$ is greater than or equal to 1,
$\Leftrightarrow$ $(B - \lambda I)$ is singular,
$\Leftrightarrow$ determinant of $(B - \lambda I)$ is zero.

Moreover, if there is a basis of eigenvectors say $\{v_1, v_2, \ldots, v_n\}$, with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then image of any vector is completely known. Suppose if $z \in \mathbb{R}^n$, then there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$z = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n, \tag{5.1}$$

so that

$$Bz = \alpha_1 Bv_1 + \alpha_2 Bv_2 + \ldots + \alpha_n Bv_n = \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \ldots + \alpha_n \lambda_n v_n.$$

In this case the matrix of linear transformation with respect to the basis $\{v_1, v_2, \ldots, v_n\}$ turns out to be diagonal with diagonal entries as $\lambda_1, \lambda_2, \ldots, \lambda_n$ and we say that the matrix B is diagonalizable.

Exercise 5.1. Show that if a matrix B is diagonalizable, with eigenvalues $|\lambda_1| \ge |\lambda_2| \ge \ldots \ge |\lambda_n|$, then the image of the unit ball $\{x : \|x\| \le 1\}$ is contained in a ball of radius $|\lambda_1|$ with center at origin.

This exercise shows the importance of eigenvalue of largest magnitude.

5.1. The Power Method

This method is useful to find the dominant eigenvalue among a collection of eigenvalues of a matrix and an eigenvector corresponding to the dominant eigenvalue. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be a set of eigenvalues of an square matrix of order $n \times n$, with corresponding eigenvectors $v_1, v_2, \ldots, v_m$, $m \le n$ such that $z = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m$, with $\alpha_1 \ne 0$ and $|\lambda_1| > |\lambda_2| \ge |\lambda_3| \ge \ldots \ge |\lambda_m|$.

Thus,

$$B^k z = \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \ldots + \alpha_m \lambda_m^k v_m. \tag{5.2}$$

Note that if u is a vector such that $\langle v_1, u \rangle \ne 0$, then $\langle z, u \rangle \ne 0$ and

$$\frac{\langle B^{k+1} z, u \rangle}{\langle B^k z, u \rangle} = \lambda_1 \frac{\alpha_1 \langle v_1, u \rangle + \alpha_2 (\lambda_2/\lambda_1)^{k+1} \langle v_2, u \rangle + \ldots + \alpha_m (\lambda_m/\lambda_1)^{k+1} \langle v_m, u \rangle}{\alpha_1 \langle v_1, u \rangle + \alpha_2 (\lambda_2/\lambda_1)^k \langle v_2, u \rangle + \ldots + \alpha_m (\lambda_m/\lambda_1)^k \langle v_m, u \rangle}. \tag{5.3}$$

So that in the limiting case

$$\lim_{k \to \infty} \frac{\langle B^{k+1} z, u \rangle}{\langle B^k z. u \rangle} = \lambda_1. \tag{5.4}$$

be diagonal with diagonal entries as $\lambda_1, \lambda_2, \ldots, \lambda_n$ and we say that the matrix B is diagonalizable.

5.1. The Power Method

This method is useful to find the dominant eigenvalue among a collection of eigenvalues of a matrix and an eigenvector corresponding to the dominant eigenvalue. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be a set of eigenvalues of an square matrix of order $n \times n$, with corresponding eigenvectors $v_1, v_2, \ldots, v_m$, $m \leq n$ such that $z = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m$, with $\alpha_1 \neq 0$ and $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_m|$.
Thus,

$$B^k z = \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \ldots + \alpha_m \lambda_m^k v_m. \tag{5.2}$$

Note that if u is a vector such that $\langle v_1, u\rangle \neq 0$, then $\langle z, u\rangle \neq 0$ and

$$\frac{\langle B^{k+1} z, u\rangle}{\langle B^k z, u\rangle} = \lambda_1 \frac{\alpha_1 \langle v_1, u\rangle + \alpha_2 (\lambda_2/\lambda_1)^{k+1} \langle v_2, u\rangle + \ldots + \alpha_m (\lambda_m/\lambda_1)^{k+1} \langle v_m, u\rangle}{\alpha_1 \langle v_1, u\rangle + \alpha_2 (\lambda_2/\lambda_1)^{k} \langle v_2, u\rangle + \ldots + \alpha_m (\lambda_m/\lambda_1)^{k} \langle v_m, u\rangle}. \tag{5.3}$$

So that in the limiting case

$$\lim_{k\to\infty} \frac{\langle B^{k+1} z, u\rangle}{\langle B^k z, u\rangle} = \lambda_1. \tag{5.4}$$

Moreover, from (5.2) $\lambda_1^{-k} B^k z = \alpha_1 v_1 + \alpha_2 (\lambda_2/\lambda_1)^k v_2 + \ldots + \alpha_m (\lambda_m/\lambda_1)^k v_m$. Thus

$$\lim_{k\to\infty} \lambda_1^{-k} B^k z = \alpha_1 v_1. \tag{5.5}$$

Thus by (5.4), we can first find the largest eigenvalue and then by (5.2) the eigenvector, corresponding to the largest eigenvalue, involved in the representation of z. Note that the eigenvector $\lambda_1 v_1$ is not necessarily of unit length.

Remark 5.1. Note that since $\|\cdot\|$ is a continuous function, from (5.5) we have $\lim_{k\to\infty} \|\lambda_1^{-k} B^k z\| = \|\alpha_1 v_1\|$. And hence $\lim_{k\to\infty} \frac{\lambda_1^{-k} B^k z}{\|\lambda_1^{-k} B^k z\|} = \frac{\alpha_1 v_1}{\|\alpha_1 v_1\|}$, or $\lim_{k\to\infty} \frac{B^k z}{\|B^k z\|} = \frac{v_1}{\|v_1\|}$. Further since B represents a continuous linear map from $\mathbb{R}^n$ to $\mathbb{R}^n$, we have $\lim_{k\to\infty} B\left(\frac{B^k z}{\|B^k z\|}\right) = B\left(\frac{v_1}{\|v_1\|}\right)$, or equivalently,

$$\lim_{k\to\infty} \frac{B^{k+1} z}{\|B^k z\|} = \lambda_1 \frac{v_1}{\|v_1\|} \tag{5.6}$$

QR Decomposition

$$\begin{aligned} a_1 &= \langle a_1, e_1\rangle e_1, \\ a_2 &= \langle a_2, e_1\rangle e_1 + \langle a_2, e_2\rangle e_2 \\ \text{In general ,} \quad e_k &= \sum_{j=1}^{k} \langle a_k, e_j\rangle e_j. \end{aligned}$$

Thus if we consider

$$Q = [e_1, e_2, \ldots, e_n]^t, \quad \text{and} \quad R = \begin{bmatrix} \langle a_1, e_1\rangle & \langle a_2, e_1\rangle & \ldots & \langle a_n, e_1\rangle \\ 0 & \langle a_2, e_2\rangle & \ldots & \langle a_n, e_2\rangle \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \langle a_n, e_2\rangle \end{bmatrix},$$

then $A = QR$.

Note that the norm used here is $\|\cdot\|_2$, which is compatible with inner product. Moreover, it can also be shown that QR decomposition of a non-singular square matrix is unique.

Problem 5.2. Find the QR decomposition of the matrix $\begin{bmatrix} -3 & -5 & -8 \\ 6 & 4 & 1 \\ -6 & 2 & 5 \end{bmatrix}$.

Solution. Note that $a_1 = [-3, 6, -6]^t$, $a_2 = [-5, 4, 2]^t$, and $a_3 = [-8, 1, 5]^t$. So that $\|a_1\| = \sqrt{9+36+36} = 9$ and $e_1 = [-1/3, 2/3, -2/3]^t$. Now $\langle a_1, e_1\rangle = 9$, $\langle a_2, e_1\rangle = 3$, $\langle a_3, e_1\rangle = 0$ so that $u_2 = [-5, 4, 2]^t - 3[-1/3, 2/3, -2/3]^t = [-4, 2, 4]^t$. Thus $e_2 = [-2/3, 1/3, 2/3]^t$ and $\langle a_2, e_2\rangle = 6$, $\langle a_3, e_2\rangle = 9$. Now $u_3 = [-8, 1, 5]^t - 0e_1 - 9[-2/3, 1/3, 2/3]^t = [-2, -2, -1]^t$ so that $e_3 = [-2/3, -2/3, -1/3]^t$ and $\langle a_3, e_3\rangle = 3$. Thus $Q = \frac{1}{3}\begin{bmatrix} -1 & -2 & -2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$ and $R = \begin{bmatrix} 9 & 3 & 0 \\ 0 & 6 & 9 \\ 0 & 0 & 3 \end{bmatrix}$.

5.3. QR Algorithm

Let A_1 be a square matrix of order n with n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|$. Decompose A_1 as $A_1 = Q_1 R_1$, where R_1 is an upper triangular matrix and Q_1 is orthogonal matrix such that $Q_1^t = Q_1^{-1}$. Consider $A_2 = R_1 Q_1 = Q_1^{-1} Q_1 R_1 Q_1 = Q_1^{-1} A_1 Q_1$. Thus A_2 is similar to A_1 and hence the set of eigenvalues of A_2 is same as the set of eigenvalues of A_1. Now if A_2 has the QR decomposition as $A_2 = Q_2 R_2$, we define $A_3 = R_3 Q_3$, which is again similar to A_2 and hence similar to A_1. Thus we find a sequence of similar matrices $\{A_n\}$.

For the convergence following theorem is stated without proof.

Theorem 5.1. *If a matrix A of order $n \times n$ has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|$ and all the principal minors of the matrix of eigenvectors of A^t are non zero, then sequence $\{A_n\}$ converges to a diagonal matrix with diagonal entries as eigenvalues.*

If λ is an eigenvalue of a square matrix B with eigenvector v, then $Bv = \lambda v$, or $\|Bv\| = |\lambda|\|v\|$ and hence

$$|\lambda| = \frac{\|Bv\|}{\|v\|} \leq \frac{\|B\| \times \|v\|}{\|v\|} = \|B\|.$$

Note that this is true for all possible matrix norms, that is, $\|\cdot\|_1, \|\cdot\|_2$ and $\|\cdot\|_\infty$. Thus

$$|\lambda| \leq \min\{\|B\|_1, \|B\|_2, \|B\|_\infty\}$$

Theorem 5.2. *(**Gershgorin's Theorem**) Let A be a square matrix B of order $n \times n$. Each eigenvalue λ of B satisfies*

$$|a_{ii} - \lambda| \leq \sum_{j=1, j\neq i}^{n} |a_{ij}|, \tag{5.7}$$

at least for some $1 \leq i \leq n$.

Remark 5.2. Since the set of eigenvalues of a square matrix A is as of its transpose A^t, one can apply the Gershgorin's theorem to A^t to conclude

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

Problem 5.4. Use Gerschgorin's theorem to find the location of eigenvalues of the matrix

$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & -2 & 1 \\ 2 & -1 & -1 \end{pmatrix}.$$

Solution. Let λ be eigenvalue of the given matrix. According to Gerschgorin's theorem the λ has to satisfy at least one of the following conditions. $|\lambda - 1| \leq 1$, $|\lambda + 2| \leq 1 + 1$ and $|\lambda + 1| \leq 2 + 1$. Thus all the eigenvalues of the matrix lie within the union of these three disks. Further if we apply Gershgorin's theorem to transpose of the given matrix, then λ should lie within the union of the disks $|\lambda - 1| \leq 1 + 2$, $|\lambda + 2| \leq 1$, and $|\lambda + 1| \leq 1 + 1$. Thus finally we conclude that all the eigenvalues should lie within the intersection of these two unions.

$$\begin{aligned} \delta x &= (A + \delta A)^{-1}(-\delta A x + \delta b) \\ &= [A(I + A^{-1}\delta A)]^{-1}(-\delta A x + \delta b) \\ &= (I + A^{-1}\delta A)^{-1} A^{-1}(-\delta A x + \delta b). \end{aligned}$$

Taking norms, dividing both sides by $\|x\|$, using part 1 of Lemma 1.7 and the triangle inequality, and assuming that δA is small enough so that $\|A^{-1}\delta A\| \leq \|A^{-1}\| \cdot \|\delta A\| < 1$, we get the desired bound:

$$\begin{aligned} \frac{\|\delta x\|}{\|x\|} &\leq \|(I + A^{-1}\delta A)^{-1}\| \cdot \|A^{-1}\| \left(\|\delta A\| + \frac{\|\delta b\|}{\|x\|} \right) \\ &\leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\delta A\|} \left(\|\delta A\| + \frac{\|\delta b\|}{\|x\|} \right) \quad \text{by Lemma 2.1} \\ &= \frac{\|A^{-1}\| \cdot \|A\|}{1 - \|A^{-1}\| \cdot \|A\| \frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|A\| \cdot \|x\|} \right) \\ &\leq \frac{\kappa(A)}{1 - \kappa(A) \frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right) \qquad (2.4) \\ & \text{since } \|b\| = \|Ax\| \leq \|A\| \cdot \|x\|. \end{aligned}$$