Numerical Analysis

By Krishna Shinde
Department of Mathematics
Modern College of Arts, Science and Commerce(Autonomous)
Shivajinagar, Pune - 5

CHAPTER 3

Eigenvalues and Eigenvectors

In this chapter, we will develop a variety of techniques for approximating the eigenvalues and eigenvectors of $n \times n$ matrix.

Definition. An eigenvalue of a matrix A is any number λ, for which the equation $Av = \lambda v$ has a nonzero solution for the vector v.

Since the equation $Av = \lambda v$ is equivalent to $(A - \lambda I)v = 0$, we see that the eigenvalue of A are those values of λ for which the matrix $A - \lambda I$ is singular; that is, those values of λ for which the $\det(A - \lambda I) = 0$ is singular.

As a function of λ, $\det(A - \lambda I)$ is a nth degree polynomial, known as characteristic polynomial of A. Counting multiplicities, and $n \times n$ matrix has precisely n eigenvalues. Furthermore, the coefficients of the characteristic polynomial are sum and product of elements of A. If A is a real matrix, then eigenvalues of A are real or occur in complex conjugate pairs. The collection eigenvalues of A is called as spectrum of the matrix.

A nonzero vector v for which $Av = \lambda v$ is called an eigenvector of the matrix A associated with the eigenvalue λ. Since v is solution to the matrix equation $(A - \lambda)v = 0$ when $a - \lambda I$ is singular, the eigenvectors are not unique. They are however determined up to a multiplicative constants. In other ward, if v is an eigenvalue associated with eigenvalue λ, the αv is also eigenvector associated with the same eigenvalue, for any nonzero constant α.

Localizing Eigenvalues

Theorem. Let A be an $n \times n$ matrix and define $r_i = \sum_{j=1, j \neq i}^{n} |a_{ij}|$ for each $i = 1, 2, ..., n$. Further, let

$$C_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$$

where $\mathbb{C}$ denote complex plane. If λ is eigenvalue of A then λ lies in one of the circle C_i.

PROOF. Let λ be an eigenvalue of A, with associated eigenvector x. Define $r_i = \sum_{j=1, j \neq i}^{n} |a_{ij}|$ for each $i = 1, 2, ..., n$. Further, let k be an index for which $|x_k| = \|x\|_\infty$. Equating the kth element in the eigenvalue relation $Ax = \lambda x$ yields

$$\sum_{j=1}^{n} a_{kj} x_j = \lambda x_k$$

or

$$(\lambda - a_{kk})x_k = \sum_{j=1}^{k-1} a_{kj}x_j - \sum_{j=k+1}^{n} a_{kj}x_j$$

Hence, upon taking the absolute value and repeatedly applying the triangle inequality,

$$\begin{aligned}
|\lambda - a_{kk}||x_k| &\leq \left|\sum_{j=1}^{k-1} a_{kj}x_j\right| - \left|\sum_{j=k+1}^{n} a_{kj}x_j\right| \\
&\leq \|x\|_\infty \left|\sum_{j=1}^{k-1} a_{kj}\right| - \|x\|_\infty \left|\sum_{j=k+1}^{n} a_{kj}x_j\right| \\
&\leq r_k \|x_k\|_\infty.
\end{aligned}$$

This follows that $|\lambda - a_{kk}| \leq r_k$ and hence $\lambda \in C_k$.

THE POWER METHOD

Let A is $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, not necessarily distinct, that satisfy the relation $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$. The eigenvalues λ_1, which is largest in magnitude, is known as the dominant eigenvalue of the matrix A. Assume that the associated eigenvectors $v_1, v_2, ..., v_n$ are linearly independent, and therefore forms a basis for $\mathbb{R}^{n+1}$. Let $x^{(0)}$ be a non-zero element of $\mathbb{R}^n$. Since the eigenvector of A forms a basis for $\mathbb{R}^n$, it follows that $x^{(0)}$ can be written as a linear combination of $v_1, v_2, ..., v_n$; that is there exist constants $\alpha_1, \alpha_2, ..., \alpha_n$ such that

$$x^{(0)} = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$$

Next, construct the sequence of vectors $\left\{x^{(m)}\right\}$ according to the rule $x^{(m)} = Ax^{(m-1)}$ for $m \geq 1$. By direct calculation we find

$$\begin{aligned}
x^{(1)} &= Ax^{(0)} \\
&= A(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) \\
&= \alpha_1(Av_1) + \alpha_2(Av_2) + ... + \alpha_n(Av_n) \\
&= \alpha_1(\lambda_1 v_1) + \alpha_2(\lambda_2 v_2) + ... + \alpha_n(\lambda_n v_n) \\
x^{(2)} &= Ax^{(1)} \\
&= A(Ax^{(0)}) \\
&= A^2(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) \\
&= \alpha_1(A^2 v_1) + \alpha_2(A^2 v_2) + ... + \alpha_n(A^2 v_n) \\
&= \alpha_1(\lambda_1^2 v_1) + \alpha_2(\lambda_2^2 v_2) + ... + \alpha_n(\lambda_n^2 v_n)
\end{aligned}$$

and, in general,

$$\begin{aligned} x^{(m)} &= Ax^{(m-1)} = A^m(x^{(0)}) \\ &= A^m(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) \\ &= \alpha_1(A^m v_1) + \alpha_2(A^m v_2) + ... + \alpha_n(A^m v_n) \\ &= \alpha_1(\lambda_1^m v_1) + \alpha_2(\lambda_2^m v_2) + ... + \alpha_n(\lambda_n^m v_n) \end{aligned}$$

In deriving these expressions we have made repeated use of $Av_j = \lambda v_j$, which follows from the fact that v_j is an eigenvector associated with the eigenvalue λ_j.

Factoring λ_1^m from the right-hand side of the equation for $x^{(m)}$ gives

$$x^{(m)} = \lambda_1^m \left[\alpha_1 v_1 + \alpha_2 \left(\frac{\lambda_2^m}{\lambda_1^m}\right) v_2 + ... + \alpha_n \left(\frac{\lambda_2^m}{\lambda_1^m}\right) v_n\right]$$

By assumption $|\lambda_j/\lambda_1| < 1$ for each j, so $|\lambda_j/\lambda_1|^m \to 0$ as $m \to \infty$. Therefore,

$$\lim_{m\to\infty} \frac{x^{(m)}}{\lambda_1^m} = \alpha_1 v_1.$$

Since any non-zero constant multiple times an eigenvector is still an eigenvector associated with the same eigenvalue. Hence the scaled sequence $\left\{x^{(m)}/\lambda_1^m\right\}$ converses to an eigenvector associated with the dominant eigenvalue provided, $\alpha_1 \neq 0$.

An approximation for the dominant eigenvalue of A can be obtained from the sequence $\left\{x^{(m)}\right\}$ as follows. Let i be an index for which $x_i^{(m-1)} \neq 0$, and consider the ration of the i element from the vector $x^{(m)}$ to the ith element from $x^{(m-1)}$

$$\frac{x_i^{(m)}}{x_i^{(m-1)}} = \frac{\lambda_1^m \left[\alpha_1 v_{1,i} + \alpha_2 \left(\frac{\lambda_2^m}{\lambda_1^m}\right) v_{2,i} + ... + \alpha_n \left(\frac{\lambda_n^m}{\lambda_1^m}\right) v_{n,i}\right]}{\lambda_1^{m-1} \left[\alpha_1 v_{1,i} + \alpha_2 \left(\frac{\lambda_2^{m-1}}{\lambda_1^{m-1}}\right) v_{2,i-1} + ... + \alpha_n \left(\frac{\lambda_n^{m-1}}{\lambda_1^{m-1}}\right) v_{n,i-1}\right]}$$

Since, $|\lambda_j/\lambda_1| < 1$ for each j, so $|\lambda_j/\lambda_1|^{m-1}, |\lambda_j/\lambda_1|^m \to 0$ as $m \to \infty$.

$$\begin{aligned} \lim_{m\to\infty} \frac{x_i^{(m)}}{x_i^{(m-1)}} &= \frac{\lambda_1^m (\alpha_1 v_{1,i})}{\lambda_1^{m-1} (\alpha_1 v_{1,i})} \\ &= \lambda_1 \end{aligned}$$

Therefore, the sequence $\left\{\frac{x_i^{(m)}}{x_i^{(m-1)}}\right\}$ converges to dominant eigenvalue λ_1.

To simplify the notations, let's introduce the vector $y^{(m)}$ to denote the result of multiplication by the matrix A; that is, $y^{(m)} = Ax^{(m-1)}$. $x^{(m)}$ is then calculated by the formula

$$x^{(m)} = \frac{y^{(m)}}{y_{p_m}^{(m)}},$$

where p_m is an integer chosen so that $|y^{(m)}_{p_m}| = \|y^{(m)}\|_\infty$. Note that p_m is an index into the vector $y^{(m)}$. Whenever there is more than one possible choice for the index p_m, we will adopt the convention of always selecting the smallest value. The vector x^m now converges specifically to the multiple of v_1 which has unit length measured in the infinity norm. As for the eigenvalue, since $x^{(m-1)}$ is approximately an eigenvector associated with λ_1, $y^{(m)} = Ax^{(m-1)} \approx \lambda_1 x^{(m-1)}$. By construction $x^{(m-1)}_{p_{m-1}} = 1$, so it follows that $y^{(m)}_{p_{m-1}}$ converges to λ_1.

Example. Find the dominant eigenvalue and corresponding eigenvector of a matrix $A = \begin{bmatrix} -2 & -2 & 3 \\ -10 & -1 & 6 \\ 10 & -2 & -9 \end{bmatrix}$ whose eigenvalues are $\lambda_1 = -12, \lambda_2 = -3$ and $\lambda_3 = 3$.

Solution. Let us start with vector $x^{(0)} = [1 \;\; 0 \;\; 0]^T \implies \|x^{(0)}\|_\infty = 1$.
Therefore, we set $p_0 = 1$(initially we consider $x^{(0)} = y^{(0)}$).
For the first iteration of the power method we compute,

$$\begin{aligned} y^{(1)} &= Ax^{(0)} \\ &= \begin{bmatrix} -2 & -2 & 3 \\ -10 & -1 & 6 \\ 10 & -2 & -9 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ -10 \\ 10 \end{bmatrix} \end{aligned}$$

from which we obtain the first approximation to dominant eigenvalue: $\lambda^{(1)} = y^{(1)}_{p_0} = y^{(1)}_1 = -2$.
Since $\|y^{(1)}\|_\infty = 10$. For our convenience of selecting the smallest index for which the magnitude of the vector element is equal to the infinity norm of the vector, we take $p_1 = 2$. Therefore, for the second iteration, we have

$$\begin{aligned} x^{(1)} &= \frac{y^{(1)}}{y^{(1)}_{p_1}} \\ &= -\frac{1}{10} \begin{bmatrix} -2 \\ -10 \\ 10 \end{bmatrix} \\ &= \begin{bmatrix} 1/5 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

The calculations for the second iteration produce the results

$$\begin{aligned} y^{(2)} &= Ax^{(1)} \\ &= \begin{bmatrix} -2 & -2 & 3 \\ -10 & -1 & 6 \\ 10 & -2 & -9 \end{bmatrix} \begin{bmatrix} 1/5 \\ 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} -27/5 \\ -9 \\ 9 \end{bmatrix} \end{aligned}$$

$$\lambda^{(2)} = y^{(2)}_{p_1} = y^{(2)}_2 = -9$$

$$p_2 = 2$$

and

$$\begin{aligned} x^{(2)} &= \frac{y^{(2)}}{y^{(2)}_{p_2}} \\ &= -\frac{1}{9} \begin{bmatrix} -27/5 \\ -9 \\ 9 \end{bmatrix} \\ &= \begin{bmatrix} 3/5 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

The third iteration then produces

$$\begin{aligned} y^{(3)} &= Ax^{(2)} \\ &= \begin{bmatrix} -2 & -2 & 3 \\ -10 & -1 & 6 \\ 10 & -2 & -9 \end{bmatrix} \begin{bmatrix} 3/5 \\ 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} -31/5 \\ -13 \\ 13 \end{bmatrix} \\ \lambda^{(3)} &= y^{(3)}_{p_2} = y^{(3)}_2 = -13, \\ p_3 &= 2 \end{aligned}$$

and

$$x^{(3)} = \frac{y^{(3)}}{y^{(3)}_{p_3}}$$

$$= -\frac{1}{13}\begin{bmatrix} -31/5 \\ -13 \\ 13 \end{bmatrix}$$

$$= \begin{bmatrix} 31/5 \\ 1 \\ -1 \end{bmatrix}$$

The following table displays the output of 11 iterations of power method

j	$x^{(j)^T}$	λ^j
0	[1.000000 0.000000 0.000000]	
1	[0.200000 1.000000 − 1.000000]	-2.00000
2	[0.600000 1.000000 − 1.000000]	-9.000000
3	[0.476923 1.000000 − 1.000000]	-13.000000
4	[0.505882 1.000000 − 1.000000]	-11.769231
5	[0.498537 1.000000 − 1.000000]	-12.058824
6	[0.500366 1.000000 − 1.000000]	-11.985366
7	[0.499908 1.000000 − 1.000000]	-12.003663
8	[0.500023 1.000000 − 1.000000]	-11.999085
9	[0.499994 1.000000 − 1.000000]	-12.000229
10	[0.500001 1.000000 − 1.000000]	-11.999943
11	[0.500000 1.000000 − 1.000000]	-12.000014

The final estimate are

$$\lambda_1 \approx -12.000014 \text{ and } v_1 \approx [0.500000 \quad 1.000000 \quad -1.000000]^T$$

Power Method for Symmetric Matrices

When a matrix A is symmetric, a slight modification to the power method provides more rapid convergence. In this method we select the initial vector $x^{(0)}$ be a non-zero element of $\mathbb{R}^n$ with $x^{(0)^T}x^{(0)} = 1$. The modified iteration schemes are as follows:

$$y^{(m)} = Ax^{(m-1)}$$

$$\lambda^{(m)} = x^{(m-1)^T}y^{(m)} \text{ and}$$

$$x^{(m)} = y^{(m)}/\sqrt{y^{(m)^T}y^{(m)}}.$$

Then $\lambda^m \to \lambda_1$ and $x^{(m)}$ converges to an associated with λ_1 that has unit length in the Euclidean norm.

Example. Find the dominant eigenvalue of 4×4 symmetric matrix

$$A = \begin{bmatrix} 5.5 & -2.5 & -2.5 & -1.5 \\ -2.5 & 5.5 & 1.5 & 2.5 \\ -2.5 & 1.5 & 5.5 & 2.5 \\ -1.5 & 2.5 & 2.5 & 5.5 \end{bmatrix},$$

whose eigenvalues are $\lambda_1 = 12, \lambda_2 = 4, \lambda_3 = 4$ and $\lambda_4 = 2$. The eigenvector associated with eigenvalue λ_1 that has unit Euclidean norm is $v_1 = [-1/2 \ \ 1/2 \ \ 1/2 \ \ 1/2]^T$.
Solution. We will start the iteration with the vector $x^{(0)} = [0.5 \ \ 0.5 \ \ 0.5 \ \ 0.5]^T$. Here

$$x^{(0)^T} x^{(0)} = (0.5)(0.5) + (0.5)(0.5) + (0.5)(0.5) + (0.5)(0.5) = 1,$$

For $m = 1$, we calculate

$$\begin{aligned} y^{(1)} &= Ax^{(0)} \\ &= \begin{bmatrix} 5.5 & -2.5 & -2.5 & -1.5 \\ -2.5 & 5.5 & 1.5 & 2.5 \\ -2.5 & 1.5 & 5.5 & 2.5 \\ -1.5 & 2.5 & 2.5 & 5.5 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix} \\ &= \begin{bmatrix} -0.5 \\ 3.5 \\ 3.5 \\ 4.5 \end{bmatrix} \\ \lambda^{(1)} &= x^{(0)^T} y^{(1)} \\ &= (0.5)(-0.5) + (0.5)(3.5) + (0.5)(3.5) + (0.5)(4.5) \\ &= 5.5 \end{aligned}$$

and

$$\begin{aligned} x^{(1)} &= \frac{y^{(1)}}{\sqrt{y^{(1)^T} y^{(1)}}} \\ &= \frac{1}{\sqrt{(-0.5)(-0.5) + (3.5)(3.5) + (3.5)(3.5) + (4.5)(4.5)}} \begin{bmatrix} -0.5 \\ 3.5 \\ 3.5 \\ 4.5 \end{bmatrix} \\ &= \begin{bmatrix} -0.074536 \\ 0.521749 \\ 0.521749 \\ 0.670820 \end{bmatrix} \end{aligned}$$

Continue on to the second iteration, we find

$$
\begin{aligned}
y^{(2)} &= Ax^{(1)} \\
&= \begin{bmatrix} 5.5 & -2.5 & -2.5 & -1.5 \\ -2.5 & 5.5 & 1.5 & 2.5 \\ -2.5 & 1.5 & 5.5 & 2.5 \\ -1.5 & 2.5 & 2.5 & 5.5 \end{bmatrix} \begin{bmatrix} -0.074536 \\ 0.521749 \\ 0.521749 \\ 0.670820 \end{bmatrix} \\
&= \begin{bmatrix} -4.024920 \\ 5.515630 \\ 5.515630 \\ 6.410060 \end{bmatrix} \\
\lambda^{(2)} &= x^{(1)^T} y^{(2)} \\
&= [-0.074536 \;\; 0.521749 \;\; 0.521749 \;\; 0.670820] \begin{bmatrix} -4.024920 \\ 5.515630 \\ 5.515630 \\ 6.410060 \end{bmatrix} \\
&= 10.355556
\end{aligned}
$$

and

$$
\begin{aligned}
x^{(2)} &= \frac{y^{(2)}}{\sqrt{y^{(2)^T} y^{(2)}}} \\
&= \frac{1}{10.86891} \begin{bmatrix} -4.024920 \\ 5.515630 \\ 5.515630 \\ 6.410060 \end{bmatrix} \\
&= \begin{bmatrix} -0.370315 \\ 0.507469 \\ 0.507469 \\ 0.589761 \end{bmatrix}
\end{aligned}
$$

The table below displays the result of 10 iterations.

j	$x^{(j)^T}$	λ^j
0	$[0.500000\ 0.500000\ 0.500000\ 0.500000]$	
1	$[-0.074536\ 0.521749\ 0.521749\ 0.670820]$	5.500000
2	$[-0.370315\ 0.507469\ 0.507469\ 0.589761]$	10.355556
3	$[-0.460013\ 0.501622\ 0.501622\ 0.533985]$	11.799850
4	$[-0.487194\ 0.500309\ 0.500309\ 0.511882]$	11.977899
5	$[-0.495812\ 0.500056\ 0.500056\ 0.504042]$	11.997556
6	$[-0.498617\ 0.500010\ 0.500010\ 0.501360]$	11.999729
7	$[-0.499541\ 0.500002\ 0.500002\ 0.500455]$	11.999970
8	$[-0.499847\ 0.500000\ 0.500000\ 0.500152]$	11.999997
9	$[-0.499949\ 0.500000\ 0.500000\ 0.500051]$	12.000000
10	$[-0.499983\ 0.500000\ 0.500000\ 0.500017]$	12.000000

THE INVERSE POWER METHOD

The power method is designed to approximate the dominant eigenvalue of a matrix. There are many instances, however, in which an eigenvalue other than dominant one is needed. To approximate the other eigenvalues inverse power method is used.

Theorem. Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ and associated eigenvector $v_1, v_2, ..., v_n$.

1. If $B = a_0 + a_1 A + a_2 A^2 + \cdots + a_m A^m = p(A)$, where p is the polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$, then the eigenvalues of B are $p(\lambda_1), p(\lambda_2), ..., p(\lambda_n)$ with associated eigenvectors $v_1, v_2, ..., v_n$.

2. If A is non-singular, then A^{-1} has eigenvalues

$$\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \frac{1}{\lambda_3}, ..., \frac{1}{\lambda_n}$$

PROOF. Part 1:

For any positive integer k,

$$\begin{aligned} A^k v_i &= A^{k-1}(Av_i) = \lambda_i A^{k-1} v_i \\ &= \lambda_i A^{k-2}(Av_i) = \lambda_i^2 A^{k-2} v_i \\ &= ... \\ &= \lambda_i^{k-1}(Av_i) = \lambda_i^k v_i. \end{aligned}$$

Now, let $B = a_0 I + a_1 A + a_2 A^2 + ... + a_m A^m = p(A)$, where p is the polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_m x^m$. Then, for each $i = 1, 2, 3, ..., n$,

$$\begin{aligned} Bv_i &= (a_0 I + a_1 A + a_2 A^2 + ... + a_m A^m) v_i \\ &= a_0 v_i + a_1 A v_i + a_2 A^2 v_i + ... + a_m A^m v_i \\ &= a_0 v_i + a_1 \lambda_i v_i + a_2 \lambda_i^2 v_i + ... + a_m \lambda_i^m v_i \\ &= (a_0 + a_1 \lambda_i + a_2 \lambda_i^2 + ... + a_m \lambda_i^m) v_i \\ &= p(\lambda_i) v_i \end{aligned}$$

Hence, the eigenvalues of B are

$$p(\lambda_1), p(\lambda_2), p(\lambda_3)..., p(\lambda_n)$$

with associated eigenvector $v_1, v_2, v_3, ..., v_n$.
Part 2:
Suppose A is non-singular. Since v_i is an eigenvector associated with the eigenvalue λ_i, it follows that

$$Av_i = \lambda_i v_i.$$

Premultiplying this by $(1/\lambda_i)A^{-1}$ yields

$$\frac{1}{\lambda_i}A^{-1}(Av_i) = \frac{1}{\lambda_i}A^{-1}(\lambda_i v_i),$$

or

$$\frac{1}{\lambda_i}v_i = A^{-1}v_i,$$

Therefore, for each $i = 1, 2, ..., n$, $1/\lambda_i$ is an eigenvalue of A^{-1}, with associated eigenvector v_i.

Method

Once again, let A be an $n \times n$ with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ and associated eigenvectors $v_1, v_2, ..., v_n$. Let q be any constant for $A - qI$ is non-singular (this will hold true for any q that is not an eigenvalue of A), and consider the matrix $B = (A - qI)^{-1}$. As a consequence of the theorem we just finished proving, the eigenvalue of B are

$$\mu_1 = \frac{1}{\lambda_1 - q}, \mu_2 = \frac{1}{\lambda_2 - q}, \mu_3 = \frac{1}{\lambda_3 - q}, ..., \mu_n = \frac{1}{\lambda_n - q}$$

with associated eigenvector $v_1, v_2, ..., v_n$.

If we apply the power method to the matrix B, the eigenvalue $\lambda^{(m)}$ will converge to the dominant eigenvalue, say μ_k. Note, however, that μ_k will be the dominant eigenvalue of B if and only if λ_k is the eigenvalue of A that is closest to the number q.

If A has an eigenvalue in the vicinity of q, we can approximate to that eigenvalue by applying the power method to the matrix B. This process is known as the inverse power method.

An implementation of the inverse power method can be obtained from code for the power method which only a few modifications. First, an extra input value, the number q, must be included in the parameter list. Second, the operation $y^{(m)} = Ax^{(m-1)}$ must be replaced by $y^{(m)} = (A-qI)^{-1}x^{(m-1)}$. In practice we solve the linear system $(A-qI)y^{(m)} = x^{(m-1)}$ for $y^{(m)}$. Since the matrix $A-qI$ can be computed once prior to the iteration loop and only the solve step(forward and backward substitution) need be performed with each iteration. Third, remember that the sequence $\lambda^{(m)}$ converges to $(\lambda_k - q)^{-1}$. To obtain an approximation to λ_k, we must compute $(1/\lambda^{(m)})+q$. The eigenvector of A and $(A-qI)^{-1}$ are the same, so no manipulation of the sequence $\{x^{(m)}\}$ is necessary.

Example. Find the eigenvalue of following matrix by inverse power method

$$A = \begin{bmatrix} 12 & 1 & 1 & 0 & 3 \\ -1 & 3 & 0 & 1 & 0 \\ 1 & 0 & -6 & 2 & 1 \\ 0 & 2 & 1 & 9 & 0 \\ 1 & 0 & 1 & 0 & -2 \end{bmatrix}$$

Solution. The Gerschgorin circles for A are plotted in the figure below. Each circle C_i corresponds to the i^{th} row from the matrix. Note that circle $C_2 = \{z \in \mathbb{C} : |z - 3| \leq 2\}$ is disjoint from the other four circles and hence is guaranteed to contain one of the five eigenvalues. From the figure it is clear that the eigenvalue in C_2 is not the dominant eigenvalue of the matrix, so power method will not locate it. However, the inverse power method can. Let's take $q = 3$, since this is the center of the Gerschgorin circle. With a starting vector of

$$[1 \;\; 1 \;\; 1 \;\; 1 \;\; 1]^T$$

The five iterations of inverse power method are listed in table below:

j	$x^{(j)^T}$	$3 + 1/\lambda^j$
0	$[1.000000 \;\; 1.000000 \;\; 1.000000 \;\; 1.000000 \;\; 1.000000]$	
1	$[-0.130952 \;\; 1.000000 \;\; -0.068452 \;\; -0.360119 \;\; 0.005952]$	4.750000
2	$[-0.087393 \;\; 1.000000 \;\; -0.083210 \;\; -0.306325 \;\; -0.033860]$	2.781069
3	$[-0.087658 \;\; 1.000000 \;\; -0.084217 \;\; -0.308045 \;\; -0.035867]$	2.779612
4	$[-0.087622 \;\; 1.000000 \;\; -0.084233 \;\; -0.307981 \;\; -0.035952]$	2.779641
5	$[-0.087621 \;\; 1.000000 \;\; -0.084234 \;\; -0.307983 \;\; -0.035955]$	2.779638

From the above table we see that $\lambda \approx 2.779638$ and the corresponding eigenvector

$$v = [-0.087621 \;\; 1.000000 \;\; -0.084234 \;\; -0.307983 \;\; -0.035955]^T$$

REDUCTION TO SYMMETRIC TRIDIAGONAL FORM

The eigenvalues of symmetric matrices are well-conditioned whereas the eigenvalues of non-symmetric matrices can be poorly conditioned because $n \times n$ symmetric matrix always possess n linearly independent eigenvectors whereas a non-symmetric matrix may not, we will restric our attention to symmetric matrices only.

To compute all the eigenvalues of a symmetric matrix, we will proceed in two stages. First, the matrix will be transformed to symmetric tridiagonal form. This stage requires a fixed, finite number of operations. In second stage we apply the iterative process of QR-algorithm on the triadiagonal matrix. The iteration generates a sequence of matrices which will converge to a diagonal matrix. The eigenvalues of diagonal matrix are, of course, just the elements along the main diagonal.

Similarity Transformation and Orthogonal Matrices

Definition. Let A be an $n \times n$ matrix and let M be an non-singular $n \times n$ matrix. The matrix $B = M^{-1}AM$ is said to be similar to A. The process of converting A to B is called as similarity transformation.

The similarity transformation does not affect any of the eigenvalue of A, we proceed as follows. The eigenvalue of B are solution of the equation $\det(B - \lambda I) = 0$; but

$$\begin{aligned}
\det(B - \lambda I) &= \det(M^{-1}AM - \lambda I) \\
&= \det\left[M^{-1}(A - \lambda)M\right] \\
&= \det(M^{-1})\det(A - \lambda I)\det(M) \\
&= \frac{1}{\det(M)}\det(A - \lambda I)\det(M) \\
&= \det(A - \lambda I)
\end{aligned}$$

Thus, $\det(B - \lambda I) = 0$ if and only if $\det(A - \lambda I) = 0$, which implies that A and B have exactly the same eigenvalues.

Definition. The $n \times n$ matrix Q is called an orthogonal matrix if $Q^{-1} = Q^T$.

Definition. A Householder matrix is any matrix of the form

$$H = I - 2ww^T$$

where w is a column vector with $w^T w = 1$.

Example. Show that Householder matrix is both symmetric and orthogonal(Exercise).

The Householder matrix are not computed explicitly, only the vector w is computed. For, once the vector w is known, the similarity transformation HAH is given by

$$\begin{aligned}
HAH &= (I - 2ww^T)A(I - 2ww^T) \\
&= A - 2ww^TA - 2Aww^T + 4ww^TAww^T,
\end{aligned}$$

which is completely determined by w. The computation of HAH can be simplified tremendously if we define $u = Aw$ and $K = w^T u = w^T Aw$. Then

$$\begin{aligned}
HAH &= A - 2ww^TA - 2Aww^T + 4ww^TAww^T \\
&= A - 2wu^T - 2uw^T + 4Kww^T \\
&= A - 2w(u^T - Kw^T) - 2(u - Kw)w^T.
\end{aligned}$$

If we now let $q = u - Kw$, then $HAH = A - 2wq^T - 2qw^T$.

The algorithm to reduce a symmetric matrix to tridiagonal form using Householder matrices involves a sequence of $n - 2$ similarity transformations as shown below diagram for the case $n = 5$.

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \xrightarrow{H_1AH_1} \begin{bmatrix} \times & \times & \times & \times & 0 \\ \times & \times & \times & \times & 0 \\ \times & \times & \times & \times & 0 \\ \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

$$\xrightarrow{H_2H_1AH_1H_2} \begin{bmatrix} \times & \times & \times & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ \times & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

$$\xrightarrow{H_3H_2H_1AH_1H_2H_3} \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

The first Householder matrix, H_1, is selected so that H_1A will have zeros in the first $n-2$ rows of the n^{th} row of A will not be affected. By symmetry, when H_1AH_1 is computed to complete the transformation, then zeros in the n^{th} column will not be changed, but zeros will appear in the first $n-2$ columns of the n^{th} row. Each subsequent Householder matrix $H_i(i = 2, 3, ..., n-2)$, is then selected so that

$$H_iH_{i-1} \cdots H_2H_1AH_1H_2 \cdots H_{i-1}H_i$$

will have zeros in the first $n-i-1$ rows of the $(n-i+1)^{th}$ column but will not affect the bottom i rows. Completing the i^{th} transformation will place zeros in the first $n-i-1$ columns of the $(n-i+1)^{th}$ row.

Determine the appropriate Householder matrix for use in each step of the above algorithm require the solution of the following fundamental problem:

Given an integer k and an $n-$dimensional column vector x, select w so that $Hx = (I - 2ww^T)x$ has zero in the first $n-k-1$ rows but leaves the last k elements in x unchanged.

To solve this problem, first note that in order for the last k elements in x to be unchanged, the last k elements in w must be zero. This guarantees that the last k rows and columns of H are identical to the identity matrix. Thus w must be of the form

$$w = [w_1 \;\; w_2 \;\; w_3 \;\; \cdots \;\; w_{n-k} \;\; 0 \cdots \;\; 0]^T.$$

Let $b = (I - 2ww^T)x$, where by construction b will have the form

$$b = [0 \;\; \cdots \;\; \alpha \;\; x_{n-k+1} \;\; \cdots \;\; x_n]^T,$$

with $n-k-1$ zero at the beginning of the vector. Since multiplication by the Householder matrix must preserve the Euclidean norm, we must have $b^Tb = x^Tx$, which implies

$$\alpha^2 = x_1^2 + x_2^2 + \cdots + x_{n-k}^2.$$

To proceed further, let's rearrange the equation defining the vector b as

$$x - 2ww^T x = b \tag{1}$$

Premultiplying equation (1) by w^T yields

$$w^T x - 2w^T ww^T x = w^T b$$

which simplifies to

$$-w^T x = \alpha w_{n-k} \tag{2}$$

upon taking into account the form of both w and b and using the fact that $w^T w = 1$. Substituting equation (2) into (1) produces

$$x + 2\alpha w_{n-k} w = b,$$

or, in component form,

$$x_i + 2\alpha w_{n-k} w_i = 0, \quad (i = 1, 2, 3, ..., n-k-1)$$

$$x_{n-k} + 2\alpha w_{n-k}^2 = \alpha.$$

From the last of these equations we see that

$$w_{n-k} = \sqrt{\tfrac{1}{2}\left(1 - \tfrac{x_{n-k}}{\alpha}\right)}.$$

To avoid cancellation error, we will choose $\text{sgn}(\alpha) = -\text{sgn}(x_{n-k})$. With w_{n-k} determine, the remaining nonzero entries in w are given by

$$w_i = -\tfrac{1}{2}\tfrac{x_i}{\alpha w_{n-k}} \quad (i = 1, 2, 3, ..., n-k-1)$$

Example. Convert the following matrix to symmetric tridiagonal form.

$$A = \begin{bmatrix} -1 & -2 & 1 & 2 \\ -2 & 3 & 0 & -2 \\ 1 & 0 & 2 & 1 \\ 2 & -2 & 1 & 4 \end{bmatrix}$$

Solution. We want to produce zeros in the first two rows of the last column of A and leave the last element in that column alone. Therefore, we are working with $k = 1$ and a vector $x = [2 \ \ -2 \ \ 1 \ \ 4]^T$. With this vector, we compute $\alpha^2 = x_1^1 + x_2^2 + x_3^2 = 2^2 + (-2)^2 + 1^2 = 9$ and $\text{sgn}(\alpha) = -\text{sgn}(x_3)$ =negative. Therefore, we choose $\alpha = -3$.

$$w_3 = \sqrt{\frac{1}{2}\left(1 - \frac{x_3}{\alpha}\right)} = \sqrt{\frac{1}{2}\left(1 - \frac{1}{-2}\right)} = \frac{\sqrt{6}}{3};$$

$$w_2 = -\frac{1}{2}\frac{x_2}{\alpha x_3} = -\frac{1}{2}\frac{-2}{-3(\sqrt{6}/3)} = -\frac{\sqrt{6}}{6} \text{ and}$$

$$w_1 = -\frac{1}{2}\frac{x_1}{\alpha x_3} = -\frac{1}{2}\frac{2}{-3(\sqrt{6}/3)} = \frac{\sqrt{6}}{6}$$

Hence, $w = [w_1 \;\; w_2 \;\; w_3 \;\; 0]^T = (\sqrt{6}/6)[1 \;\; -1 \;\; 2 \;\; 0]^T$. Next we compute

$$\begin{aligned} u &= Aw = \begin{bmatrix} -1 & -2 & 1 & 2 \\ -2 & 3 & 0 & -2 \\ 1 & 0 & 2 & 1 \\ 2 & -2 & 1 & 4 \end{bmatrix} (\sqrt{6}/6)[1 \;\; -1 \;\; 2 \;\; 0]^T = (\sqrt{6}/6)[3 \;\; -5 \;\; 5 \;\; 6]^T; \\ K &= w^T u = (\sqrt{6}/6)[1 \;\; -1 \;\; 2 \;\; 0](\sqrt{6}/6)[3 \;\; -5 \;\; 5 \;\; 6]^T = 3; \text{ and} \\ q &= u - Kw = (\sqrt{6}/6)[3 \;\; -5 \;\; 5 \;\; 6]^T - 3(\sqrt{6}/6)[1 \;\; -1 \;\; 2 \;\; 0] = (\sqrt{6}/6)[0 \;\; -2 \;\; -1 \;\; 6]^T \end{aligned}$$

Therefore,

$$\begin{aligned} H_1AH_1 &= A - 2wq^T - 2qw^T \\ &= \begin{bmatrix} -1 & -2 & 1 & 2 \\ -2 & 3 & 0 & -2 \\ 1 & 0 & 2 & 1 \\ 2 & -2 & 1 & 4 \end{bmatrix} - 2\frac{\sqrt{6}}{6}\begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \frac{\sqrt{6}}{6}[0 \;\; -2 \;\; -1 \;\; 6] \\ &\quad -2\frac{\sqrt{6}}{6}\begin{bmatrix} 0 \\ -2 \\ -1 \\ 6 \end{bmatrix} \frac{\sqrt{6}}{6}[1 \;\; -1 \;\; 2 \;\; 0] \\ &= \begin{bmatrix} -1 & -4/3 & 4/3 & 0 \\ -4/3 & 5/3 & 1 & 0 \\ 4/3 & 1 & 10/3 & -3 \\ 0 & 0 & -3 & 4 \end{bmatrix} \end{aligned}$$

For the second step of reduction, we want to produce a zero in the first row of the third column of H_1AH_1 and leave the last two elements in that column alone. Therefore, we are working with $k = 2$ and the vector $x = [4/3 \;\; 1 \;\; 10/3 \;\; -3]^T$. With this vector, we compute $\alpha^2 = 25/9$ and since $\text{sgn}(x_2)$ is positive, we choose $\alpha = -5/3$. It then follows that

$$\begin{aligned} w_2 &= \sqrt{\frac{1}{2}\left(1 - \frac{1}{-5/3}\right)} = \frac{2\sqrt{5}}{5} \\ w_1 &= -\frac{1}{2}\frac{4/3}{(-5/3)(3\sqrt{5}/5)} = \frac{\sqrt{5}}{5} \end{aligned}$$

Hence, $w = (\sqrt{5}/5)[1 \;\; 2 \;\; 0 \;\; 0]^T$. Next, we compute

$$\begin{aligned} u &= Aw = (\sqrt{5}/5)[-11/3 \;\; 2 \;\; 10/3 \;\; 0]^T; \\ K &= w^T u = 1/15; \text{ and} \\ q &= u - Kw = (\sqrt{5}/5)\left[-\tfrac{56}{15} \;\; \tfrac{28}{15} \;\; \tfrac{10}{3} \;\; 0\right]^T \end{aligned}$$

Therefore,

$$\begin{aligned}
H_2H_1AH_1H_2 &= H_1AH_1 - 2wq^T - 2qw^T \\
&= \begin{bmatrix} -1 & -4/3 & 4/3 & 0 \\ -4/3 & 5/3 & 1 & 0 \\ 4/3 & 1 & 10/3 & -3 \\ 0 & 0 & -3 & 4 \end{bmatrix} - 2\frac{2}{5}\begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}\begin{bmatrix} -\frac{56}{15} & \frac{28}{15} & \frac{10}{3} & 0 \end{bmatrix} \\
&\quad -\frac{2}{5}\begin{bmatrix} -56/15 \\ 28/15 \\ 10/3 \\ 0 \end{bmatrix}[1 \quad 2 \quad 0 \quad 0] \\
&= \begin{bmatrix} 149/75 & 68/75 & 0 & 0 \\ 68/75 & -33/25 & -5/3 & 0 \\ 0 & -5/3 & 10/3 & -3 \\ 0 & 0 & -3 & 4 \end{bmatrix}
\end{aligned}$$

♣♣♣