CHAPTER 4 (6 LECTURES)
ITERATIVE TECHNIQUES IN MATRIX ALGEBRA

1. Introduction

In this chapter we will study iterative techniques to solve linear systems. An initial approximation (or approximations) will be found, and new approximations are then determined based on how well the previous approximations satisfied the equation. The objective is to find a way to minimize the difference between the approximations and the exact solution. To discuss iterative methods for solving linear systems, we first need to determine a way to measure the distance between n-dimensional column vectors. This will permit us to determine whether a sequence of vectors converges to a solution of the system. In actuality, this measure is also needed when the solution is obtained by the direct methods presented in Chapter 3. Those methods required a large number of arithmetic operations, and using finite-digit arithmetic leads only to an approximation to an actual solution of the system. We end the chapter by presenting a way to find eigenvalue (dominant) and associated eigenvector. Dominant eigenvalue plays an important role in convergence of any iterative method.

1.1. **Norms of Vectors and Matrices.**

1.2. **Vector Norms.** Let $\mathbb{R}^n$ denote the set of all $n-$dimensional column vectors with real-number components. To define a distance in $\mathbb{R}^n$ we use the notion of a norm, which is the generalization of the absolute value on $\mathbb{R}$, the set of real numbers.

Definition 1.1. *A vector norm on $\mathbb{R}^n$ is a function, $\|\cdot\|$, from $\mathbb{R}^n$ into $\mathbb{R}$ with the following properties:*

(1) $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$, for all $x \in \mathbb{R}^n$.
(2) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.
(3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathbb{R}^n$ (triangle inequality).

Definition 1.2. *The l_2 and l_∞ norms for the vector $x = (x_1, x_2, \ldots, x_n)^t$ are defined by*

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \quad \text{and} \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Note that each of these norms reduces to the absolute value in the case $n = 1$.
The l_2 norm is called the Euclidean norm of the vector x because it represents the usual notion of distance from the origin in case x is in $\mathbb{R}^1 = \mathbb{R}$, $\mathbb{R}^2$, or $\mathbb{R}^3$. For example, the l_2 norm of the vector $x = (x_1, x_2, x_3)^t$ gives the length of the straight line joining the points $(0, 0, 0)$ and (x_1, x_2, x_3).

Example 1. *Determine the l_2 norm and the l_∞ norm of the vector $x = (-1, 1, -2)^t$.*

Sol. The vector $x = (-1, 1, -2)^t$ in $\mathbb{R}^3$ has norms

$$\|x\|_2 = \sqrt{(-1)^2 + (1)^2 + (-2)^2} = \sqrt{6}$$

and

$$\|x\|_\infty = \max\{|-1|, |1|, |-2|\} = 2$$

Definition 1.3 (Distance between vectors in $\mathbb{R}^n$)**.** *If $x = (x_1, x_2, \cdots, x_n)^t$ and $y = (y_1, y_2, \cdots, y_n)^t$ are vectors in $\mathbb{R}^n$, the l_2 and l_∞ distances between x and y are defined by*

$$\|x - y\|_2 = \left\{ \sum_{i=1}^{n} (x_i - y_i) \right\}^{1/2},$$

$$\|x - y\|_\infty = \max_{1 \leq i \leq n} |x_i - y_i|.$$

Definition 1.4 (Matrix Norm)**.** *A matrix norm on the set of all $n \times n$ matrices is a real-valued function, $\|\cdot\|$, defined on this set, satisfying for all $n \times n$ matrices A and B and all real numbers α :*

(1) $\|A\| \geq 0$ *and* $\|A\| = 0$ *if and only if* A *is* O*, the matrix with all* 0 *entries.*
(2) $\|\alpha A\| = |\alpha| \|A\|$.
(3) $\|A + B\| \leq \|A\| + \|B\|$.

Additionally, in the case of square matrices, some (but not all) matrix norms satisfy the following condition:

$$\|AB\| \leq \|A\| \, \|B\|.$$

If $\|\cdot\|$ *is a vector norm on* $\mathbb{R}^n$*, then*

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

is a matrix norm.

The matrix norms we will consider have the forms

$$\|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty,$$

and

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2.$$

Theorem 1.5. *If* $A = (a_{ij})$ *is an* $n \times n$ *matrix, then*

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Example 2. *Determine* $\|A\|_\infty$ *for the matrix*

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{bmatrix} \tag{1.1}$$

Sol. We have

$$\sum_{j=1}^{3} |a_{1j}| = |1| + |2| + |-1| = 4, \quad \sum_{j=1}^{3} |a_{2j}| = |0| + |3| + |-1| = 4$$

and

$$\sum_{j=1}^{3} |a_{3j}| = |5| + |-1| + |1| = 7.$$

So above theorem implies that $\|A\|_\infty = \max\{4, 4, 7\} = 7$.

Definition 1.6 (Eigenvalue and eigenvector)**.** *Let* A *be a square matrix then number* λ *is called an eigenvalue of* A *if there exists a nonzero vector* x *such that* $Ax = \lambda x$*. Here* x *is called the corresponding eigenvector.*

Definition 1.7 (Characteristic polynomial)**.** *Characteristic polynomial is defined as*

$$P(\lambda) = \det(A - \lambda I).$$

λ *is an eigenvalue of matrix* A *if and only if* λ *is a root of the characteristic polynomial, i.e.,* $P(\lambda) = 0$.

Definition 1.8 (Spectral Radius)**.** *The spectral radius* $\rho(A)$ *of a matrix* A *is defined by*

$$\rho(A) = \max |\lambda|, \text{ where } \lambda \text{ is an eigenvalue of } A.$$

Theorem 1.9. *If* A *is an* $n \times n$ *matrix, then*

$$\|A\|_2 = \left[\rho(A^T A)\right]^{1/2}.$$

The proof of this theorem requires more information concerning eigenvalues. We illustrate the procedure by taking an example.

Example 3. *Determine the* l_2 *norm of*

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

Sol. We calculate $\rho(A^TA)$, therefore we need to find the eigenvalues of A^TA.

$$\begin{aligned} A^TA &= \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & 4 \\ -1 & 4 & 5 \end{bmatrix}. \end{aligned}$$

Eigenvalues satisfy the following characteristics equation

$$\begin{aligned} 0 &= \det(A^TA - \lambda I) \\ &= \det \begin{bmatrix} 3-\lambda & 2 & -1 \\ 2 & 6-\lambda & 4 \\ -1 & 4 & 5-\lambda \end{bmatrix} \\ &= -\lambda^3 + 14\lambda^2 - 42\lambda = -\lambda(\lambda^2 - 14\lambda + 42) \\ \implies \lambda &= 0, 7 \pm \sqrt{7}. \end{aligned}$$

Therefore $\rho(A^t\ A) = 7 + \sqrt{7}$. Thus

$$\|A\|_2 = \left[\rho(A^T\ A)\right]^{1/2} = \sqrt{7+\sqrt{7}} \approx 3.106.$$

1.3. **Convergent Matrices.** In studying iterative matrix techniques, it is of particular importance to know when powers of a matrix become small (that is, when all the entries approach zero). Matrices of this type are called convergent.

Definition 1.10. *An $n \times n$ matrix A convergent if*

$$\lim_{k\to\infty} A^k = 0.$$

Example 4. *Show that matrix*

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

is a convergent matrix.

Sol. Computing powers of A, we obtain:

$$A^2 = \begin{bmatrix} \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}, A^3 = \begin{bmatrix} \frac{1}{8} & 0 \\ \frac{3}{16} & \frac{1}{8} \end{bmatrix},$$

and, in general,

$$A^k = \begin{bmatrix} \frac{1}{2^k} & 0 \\ \frac{k}{2^{k+1}} & \frac{1}{2^k} \end{bmatrix}.$$

So A is a convergent matrix because

$$\lim_{k\to\infty} \frac{1}{2^k} = 0,\ \lim_{k\to\infty} \frac{k}{2^{k+1}} = 0.$$

$$\therefore \lim_{k\to\infty} A^k = 0,$$

which implies matrix A is convergent.
Note that the convergent matrix A in this Example has spectral radius $\rho(A) = \frac{1}{2}$, because $\frac{1}{2}$ is the only eigenvalue of A. This illustrates an important connection that exists between the spectral radius of a matrix and the convergence of the matrix, as detailed in the following result.

Theorem 1.11. *The following statements are equivalent.*
(i) A is a convergent matrix.
(ii) $\lim_{k\to\infty} \|A^k\| = 0$, for all natural norms.
(iii) $\rho(A) < 1$.
(iv) $\lim_{k\to\infty} A^k x = 0$, for every x.

The proof of this theorem can be found in advanced texts of numerical analysis.

2. Iterative Methods

The linear system $Ax = b$ may have a large order. For such systems Gauss elimination is often too expensive in either computation time or computer memory requirements or both.
In an iterative method, a sequence of progressively iterates is produced to approximate the solution.
Jacobi and Gauss-Seidel Method: We start with an example. Let us consider a system of equations

$$\begin{aligned} 9x_1 + x_2 + x_3 &= 10 \\ 2x_1 + 10x_2 + 3x_3 &= 19 \\ 3x_1 + 4x_2 + 11x_3 &= 0. \end{aligned}$$

One class of iterative method for solving this system as follows.
We write

$$x_1 = \frac{1}{9}(10 - x_2 - x_3)$$
$$x_2 = \frac{1}{10}(19 - 2x_1 - 3x_3)$$
$$x_3 = \frac{1}{11}(0 - 3x_1 - 4x_2).$$

Let $x^{(0)} = [x_1^{(0)}\ x_2^{(0)}\ x_3^{(0)}]$ be an initial approximation of solution x. Then define an iteration of sequence

$$x_1^{(k+1)} = \frac{1}{9}(10 - x_2^{(k)} - x_3^{(k)})$$
$$x_2^{(k+1)} = \frac{1}{10}(19 - 2x_1^{(k)} - 3x_3^{(k)})$$
$$x_3^{(k+1)} = \frac{1}{11}(0 - 3x_1^{(k)} - 4x_2^{(k)}),\ k = 0, 1, 2, \ldots.$$

This is called Jacobi or method of simultaneous replacements. The method is named after German mathematician Carl Gustav Jacob Jacobi.
We start with [0 0 0] and obtain

$$x_1^{(1)} = 1.1111,\ x_2^{(1)} = 1.900,\ x_3^{(1)} = 0.0.$$
$$x_1^{(2)} = 0.9000,\ x_2^{(2)} = 1.6778,\ x_3^{(2)} = -0.9939$$

etc.
An another approach to solve the same system will be following.

$$x_1^{(k+1)} = \frac{1}{9}(10 - x_2^{(k)} - x_3^{(k)})$$
$$x_2^{(k+1)} = \frac{1}{10}(19 - 2x_1^{(k+1)} - 3x_3^{(k)})$$
$$x_3^{(k+1)} = \frac{1}{11}(0 - 3x_1^{(k+1)} - 4x_2^{(k+1)}),\ k = 0, 1, 2, \ldots.$$

This method is called Gauss-Seidel or method of successive replacements. It is named after the German mathematicians Carl Friedrich Gauss and Philipp Ludwig von Seidel. Starting with [0 0 0], we obtain

$$x_1^{(1)} = 1.1111,\ x_2^{(1)} = 1.6778,\ x_3^{(1)} = -0.9131.$$
$$x_1^{(2)} = 1.0262,\ x_2^{(2)} = 1.9687,\ x_3^{(2)} = -0.9588.$$

General Approach: We consider a system $Ax = b$ of order n and for $i = 1, 2, \cdots, n$, we write the i-th equation

$$\begin{aligned} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ii}x_i + \cdots + a_{in}x_n &= b_i \\ \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j\right) + a_{ii}x_i &= b_i. \end{aligned}$$

The Jacobi iterative method is obtained by solving the i-th equation for x_i to obtain (provided $a_{ii} \neq 0$)

$$x_i = \frac{1}{a_{ii}} \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} -a_{ij} x_j + b_i \right).$$

For each $k \geq 0$, generate the components $x_i^{(k+1)}$ from $x_i^{(k)}$ as

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} -a_{ij} x_j^{(k)} + b_i \right).$$

Now we write the above iterative scheme is matrix form. To write matrix form, we take

$$\begin{aligned} a_{ii} x_i^{(k+1)} &= \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} -a_{ij} x_j^{(k)} + b_i \right). \\ D x^{(k+1)} &= -(L+U) x^{(k)} + b, \end{aligned}$$

where D, L and U are diagonal, strictly lower triangle and upper triangle matrices, respectively and $A = L + D + U$. Here $b = [b_1 \ b_2 \ \cdots \ b_n]^t$.
If D^{-1} exists, then the Jacobi iterative scheme is

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

We write $T_j = -D^{-1}(L+U)$ and $B = D^{-1}b$ to obtain

$$x^{(k+1)} = T_j x^{(k)} + B.$$

Matrix T_j is called the iteration matrix.
For Gauss-Seidel, we write the i-th equation as

$$\begin{aligned} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{i,i-1}x_{i-1} + a_{ii}x_i + a_{i,i+1}x_{i+1} + \cdots + a_{in}x_n &= b_i \\ \sum_{j=1}^{i-1} a_{ij}x_j + a_{ii}x_i + \sum_{j=i+1}^{n} a_{ij}x_j &= b_i. \end{aligned}$$

$$\therefore \ x_i = \frac{1}{a_{ii}} \left[-\sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j + b_i \right].$$

The iterative scheme is given for each $k \geq 0$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left[-\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i \right].$$

In matrix form

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

where D, L and U are diagonal, strictly lower triangle and upper triangle matrices, respectively.

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

$$x^{(k+1)} = T_g \ x^{(k)} + B, \quad k = 0, 1, 2, \cdots.$$

Here $T_g = -(D+L)^{-1}U$ and this matrix is called iteration matrix and $B = (D+L)^{-1}b$.
Stopping Criteria: Since these techniques are iterative so we require a stopping criteria. Let ε is accuracy desired then we can use the following

$$\frac{\|x^{(k)} - x^{(k-1)}\|_\infty}{\|x^{(k)}\|_\infty} < \varepsilon.$$

Example 5. *Use the Gauss-Seidel method to approximate the solution of the following system:*

$$\begin{aligned} 4x_1 + x_2 - x_3 &= 3 \\ 2x_1 + 7x_2 + x_3 &= 19 \\ x_1 - 3x_2 + 12x_3 &= 31. \end{aligned}$$

Continue the iterations until $\dfrac{\|x^{(k+1)} - x^{(k)}\|_\infty}{\|x^{(k+1)}\|} < 0.001.$

Sol. To begin, write the system in the fixed point form

$$\begin{aligned} x_1 &= \frac{1}{4}(3 - x_2 + x_3) \\ x_2 &= \frac{1}{7}(19 - 2x_1 - x_3) \\ x_3 &= \frac{1}{12}(31 - x_1 + 3x_2) \end{aligned}$$

Iterations are given by

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{4}(3 - x_2^{(k)} + x_3^{(k)}) \\ x_2^{(k+1)} &= \frac{1}{7}(19 - 2x_1^{(k+1)} - x_3^{(k)}) \\ x_3^{(k+1)} &= \frac{1}{12}(31 - x_1^{(k+1)} + 3x_2^{(k+1)}), \quad k = 0, 1, 2, \cdots \end{aligned}$$

Start with a random vector $x^{(0)} = [0, 0, 0]^t$, the first approximation is

$$x_1^{(1)} = 0.7500$$

$$x_2^{(1)} = 2.5000$$

$$x_3^{(1)} = 3.1458.$$

Thus

$$\begin{aligned} x^{(1)} &= [0.7500, 2.5000, 3.1458]^t, & \frac{\|x^{(1)} - x^{(1)}\|_\infty}{\|x^{(1)}\|} &= 1 \\ x^{(2)} &= [0.9115, 2.0045, 3.0085]^t, & \frac{\|x^{(2)} - x^{(1)}\|_\infty}{\|x^{(2)}\|} &= 0.16470 \\ x^{(3)} &= [1.0010, 1.9985, 2.9995]^t, & \frac{\|x^{(3)} - x^{(2)}\|_\infty}{\|x^{(3)}\|} &= 0.29830 \\ x^{(4)} &= [1.000, 2.000, 3.000]^t, & \frac{\|x^{(4)} - x^{(3)}\|_\infty}{\|x^{(4)}\|} &= 0.0005 < 0.001. \end{aligned}$$

2.1. **Convergence analysis of iterative methods.** To study the convergence of general iteration techniques, we need to analyze the formula

$$x^{(k+1)} = Tx^{(k)} + B, \quad \text{for each } k = 0, 1, \cdots,$$

where $x^{(0)}$ is arbitrary. The next lemma and Theorem provide the key for this study.

Lemma 2.1. *If the spectral radius* $\rho(T) < 1$*, then* $(I - T)^{-1}$ *exists, and*

$$(I - T)^{-1} = I + T + T^2 + \cdots = \sum_{k=0}^{\infty} T^k.$$

Proof. Let x be an eigenvector corresponds to eigenvalue λ then by definition

$$\begin{aligned} Tx &= \lambda x \\ T^2x &= T(\lambda x) = \lambda Tx = \lambda^2 x \\ &\vdots \\ T^kx &= \lambda^k x. \end{aligned}$$

Since $\rho(T) < 1$, it follows that (T is convergent matrix)

$$\lim_{k\to\infty} T^k = 0.$$

Since above limit is zero, the matrix series

$$I + T + T^2 + \cdots + T^k + \cdots$$

is convergent. Now by multiplying the matrix $(I - T)$ to this series, we obtain

$$(I - T)(I + T + T^2 + \cdots + T^k + \cdots) = I.$$

Thus

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k.$$

■

Theorem 2.2 (Necessary and sufficient condition)**.** *A necessary and sufficient condition for the convergence of an iterative method is that the eigenvalue of iteration matrix T satisfy the inequality*

$$\rho(T) < 1.$$

Proof. Let

$$\rho(T) < 1.$$

The sequence of vector $x^{(k)}$ by iterative method (Gauss-Seidel) are given by

$$\begin{aligned} x^{(1)} &= Tx^{(0)} + B \\ x^{(2)} &= Tx^{(1)} + B = T(Tx^{(0)} + B) + B = T^2x^{(0)} + (T + I)B \\ &\vdots \\ x^{(k)} &= T^kx^{(0)} + (T^{k-1} + T^{k-2} + ... + T + I)B. \end{aligned}$$

Since $\rho(T) < 1$, this implies

$$\lim_{k\to\infty} T^kx^{(0)} = 0$$

Therefore

$$\lim_{k\to\infty} x^{(k)} = (I - T)^{-1}B \quad \text{as } k \to \infty.$$

Therefore, $x^{(k)}$ converges to unique solution $x = Tx + B$.
Conversely, assume that the sequence $x^{(k)}$ converges to x. Now

$$\begin{aligned} x - x^{(k)} &= Tx + B - Tx^{(k-1)} - B = T(x - x^{(k-1)}) \\ &= T^2(x - x^{(k-2)}) \\ &\vdots \\ &= T^k(x - x^{(0)}). \\ \implies \lim_{k\to\infty} T^k(x - x^{(0)}) &= \lim_{k\to\infty} (x - x^{(k)}) \\ &= x - \lim_{k\to\infty} x^{(k)} \\ &= x - x = 0. \end{aligned}$$

Thus matrix T is convergent and hence $\rho(T) < 1$. ■

Theorem 2.3. *If A is strictly diagonally dominant in $Ax = b$, then iterative method always converges for any initial starting vector.*

Proof. We assume that A is strictly diagonally dominant, hence $a_{ii} \neq 0$ and

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|, \quad i = 1, 2, \cdots, n$$

Gauss-Seidel iterations are given by

$$x^{(k+1)} = -D^{-1}(L+U)x^{(k)} + (D+L)^{-1}b$$
$$x^{(k+1)} = T_j x^{(k)} + B.$$

Method is convergent iff $\rho(T_j) < 1$.
Now

$$\|T_j\|_\infty = \| - (D+L)^{-1}U\|_\infty \leq \| - (D+L)^{-1}\|_\infty \, \|U\|_\infty = \frac{\|(D+L)^{-1}\|_\infty}{|\max a_{ii}|} < 1.$$

This shows the convergence condition for Jacobi method.

Further we prove the convergence of Gauss-Seidel method. Gauss-Seidel iterations are given by

$$x^{(k+1)} = -(D+L)^{-1}Ux^{(k)} + (D+L)^{-1}b$$
$$x^{(k+1)} = T_g x^{(k)} + B.$$

Let λ be an eigenvalue of matrix A and x be an eigenvector then

$$T_g x = \lambda x$$
$$-(D+L)^{-1}Ux = \lambda x$$
$$-Ux = \lambda(D+L)x$$
$$-\sum_{j=i+1}^{n} a_{ij} = \lambda[\sum_{j=1}^{i} a_{ij}x_j], \quad i = 1, 2, \ldots, n$$
$$-\sum_{j=i+1}^{n} a_{ij} = \lambda a_{ii}x_i + \lambda \sum_{j=1}^{i-1} a_{ij}x_j$$
$$\lambda a_{ii}x_i = -\lambda \sum_{j=1}^{i-1} a_{ij}x_j - \lambda \sum_{j=i+1}^{n} a_{ij}x_j$$
$$|\lambda a_{ii}x_i| \leq |\lambda| \sum_{j=1}^{i-1} |a_{ij}| \, |x_j| + |\lambda| \sum_{j=i+1}^{n} |a_{ij}| \, |x_j|$$

Since x is an eigenvector, $x \neq 0$, so we can take norm $||x||_\infty = 1$.
Hence

$$|\lambda| \left(|a_{ii}| - \sum_{j=1}^{i-1} |a_{ij}| \right) \leq \sum_{j=i+1}^{n} |a_{ij}|$$
$$\implies |\lambda| \leq \frac{\sum_{j=i+1}^{n} |a_{ij}|}{|a_{ii}| - \sum_{j=1}^{i-1} |a_{ij}|} \leq \frac{\sum_{j=i+1}^{n} |a_{ij}|}{\sum_{j=i+1}^{n} |a_{ij}|} = 1$$

which implies spectral radius $\rho(T) < 1$.
This implies Gauss-Seidel is convergent. ■

Example 6. *Consider the linear system*

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= -1, \\ 3x_1 + 2x_2 + 2x_3 &= 1, \\ x_1 + 2x_2 + 2x_3 &= 1. \end{aligned}$$

(1) *Compute* $\rho(T_j)$.

(2) *Compute* $\rho(T_g)$.
(3) *Comment.*

Sol. We write $A = L + D + U$.
For the given system of equations the coefficient matrix A is

$$\begin{bmatrix} 2 & 3 & 1 \\ 3 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

The given matrix A not strictly diagonally dominant as

$$|a_{11}| = 2 \not> |a_{12}| + |a_{13}| = 3$$
$$|a_{22}| = 2 \not> |a_{21}| + |a_{23}| = 5$$
$$|a_{33}| = 2 \not> |a_{31}| + |a_{32}| = 3.$$

Hence we are not sure whether the iterative methods converge or not.
Further any iterative method converges to unique solution for the system $Ax = b$ if and only if spectral radius $\rho(T) < 1$, where T is the iteration matrix.
Here

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 0 & 3 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

(a) Iteration matrix of Jacobi method is given by

$$\begin{aligned} T_j &= -D^{-1}(L+U) \\ &= -\begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0 & 3 & 1 \\ 3 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} \\ &= -\begin{bmatrix} 0 & 1.5 & 0.5 \\ 1.5 & 0 & 1 \\ 0.5 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1.5 & -0.5 \\ -1.5 & 0 & -1 \\ -0.5 & -1 & 0 \end{bmatrix}. \end{aligned}$$

The spectral radius $\rho(T_j)$ of matrix T_j is defined by

$$\rho(T_j) = \max |\lambda|, \text{ where } \lambda \text{ is an eigenvalue of } T_j.$$

Eigenvalues of T_j are

$$\lambda = -2.05655, 0.45559, 1.60096.$$

Thus, $\rho(T_j) = 2.05655 > 1$.

(b) Iteration matrix of Gauss-Seidel method is given by

$$\begin{aligned} T_g &= -(D+L)^{-1}U \\ &= -\begin{bmatrix} 0.5 & 0 & 0 \\ -0.75 & 0.5 & 0 \\ 0.5 & -0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 0 & 3 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1.5 & -0.5 \\ 0 & 2.25 & -0.25 \\ 0 & -1.5 & 0.5 \end{bmatrix}. \end{aligned}$$

The eigenvalues of T_g are $0, 2.443, 0.307$.
Hence $\rho(T_g) = 2.443 > 1$.
(c) As spectral radius for both Jacobi and Gauss-Seidel is greater than one, so both the methods will not work.

Example 7. *Show that the Gauss-Seidel method does not converge for the following system of equations:*

$$\begin{aligned} x_1 + 2x_2 - 2x_3 &= 7 \\ x_1 + x_2 + x_3 &= 2 \\ 2x_1 + 2x_2 + x_3 &= 5. \end{aligned}$$

Sol. For the given system of equations the coefficient matrix A is

$$\begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}$$

Gauss Seidel method converges to unique solution for the system $Ax = b$ if and only if spectral radius $\rho(T_g) < 1$, where iteration matrix $T_g = -(D + L)^{-1}U$.
Here

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 2 & 0 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 0 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Here $\det(D + L) = 1$ and hence $(D + L)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$

Thus

$$T_g = -\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{bmatrix}.$$

As T_g is an upper triangular matrix, so diagonal entries are eigen-values, i.e. $0, 2, 2$.
Also we can compute eigen-values directly by $\det(T - \lambda I) = \lambda(\lambda - 2)^2 = 0 \implies \lambda = 0, 2, 2$.
Thus $\rho(T) = 2 > 1$, hence Gauss-Seidel method does not converges to the unique solution.

3. The SOR method

We observed that the convergence of an iterative technique depends on the spectral radius of the matrix associated with the method. One way to select a procedure to accelerate convergence is to choose a method whose associated matrix has minimal spectral radius. These techniques are known as Successive Over-Relaxation (SOR). The SOR method is devised by applying extrapolation to the Gauss-Seidel method. This extrapolation takes the form of a weighted average between the previous iterate and the computed Gauss-Seidel iterate successively for each component. We multiply with a weight ω and to calculate $x_i^{(k+1)}$, we modify the Gauss-Seidel procedure to

$$\begin{aligned} x_i^{(k+1)} &= x_i^{(k)} + \omega(x_i^{(k+1)} - x_i^{(k)}) \\ x_i^{(k+1)} &= (1 - \omega)x_i^{(k)} + \omega\ x_i^{(k+1)}. \end{aligned}$$

The last term is calculated by Gauss-Seidel and we write

$$x_i^{(k+1)} = (1 - \omega)x_i^k + \frac{\omega}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right].$$

The choice of relaxation factor ω is not necessarily easy, and depends upon the properties of the coefficient matrix. If A is a symmetric and positive definite matrix and $0 < \omega < 2$, then the SOR method converges for any choice of initial approximate vector $x^{(0)}$.
Important Note: If a matrix A is symmetric, it is positive definite if and only if all its leading principle submatrices (minors) has a positive determinant.

Example 8. *Consider a linear system $Ax = b$, where*

$$A = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}, \ b = \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix}$$

a. Check, that the SOR method with value $\omega = 1.25$ of the relaxation parameter can be used to solve this system.
b. Compute the first iteration by the SOR method starting at the point $x^{(0)} = (0, 0, 0)^t$.

Sol. a. Let us verify the sufficient condition for using the SOR method. We have to check, if matrix A is symmetric, positive definite: A is symmetric as $A = A^T$, so let us check positive definitness:

$$\det(3) = 3 > 0,\ \det\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} = 8 > 0,\ \det(A) = 20 > 0.$$

All leading principal minors are positive and so the matrix A is positive definite. We know, that for symmetric positive definite matrices the SOR method converges for values of the relaxation parameter ω from the interval $0 < \omega < 2$.
Therefore the SOR method with value $\omega = 1.25$ can be used to solve this system.
b. The iterations of the SOR method are easier to compute by elements than in the vector form: Write the system as equations and write down the equations for the Gauss-Seidel iterations

$$\begin{aligned} x_1^{(k+1)} &= (-1 + x_2^{(k)} - x_3^{(k)})/3 \\ x_2^{(k+1)} &= (7 + x_1^{(k+1)} + x_3^{(k)})/3 \\ x_3^{(k+1)} &= (-7 - x_1^{(k+1)} + x_2^{(k+1)})/3. \end{aligned}$$

Now multiply the right hand side by the parameter ω and add to it the vector $x^{(k)}$ from the previous iteration multiplied by the factor of $(1 - \omega)$:

$$\begin{aligned} x_1^{(k+1)} &= (1-\omega)x_1^{(k)} + \omega(-1 + x_2^{(k)} - x_3^{(k)})/3 \\ x_2^{(k+1)} &= (1-\omega)x_2^{(k)} + \omega(7 + x_1^{(k+1)} + x_3^{(k)})/3 \\ x_3^{(k+1)} &= (1-\omega)x_3^{(k)} + \omega(-7 - x_1^{(k+1)} + x_2^{(k+1)})/3. \end{aligned}$$

For $k = 0$:

$$x_1^{(1)} = (1 - 1.25) \cdot 0 + 1.25 \cdot (-1 + 0 - 0)/3 = -0.41667$$

$$x_2^{(1)} = (1 - 1.25) \cdot 0 + 1.25 \cdot (7 - 0.41667 + 0)/3 = 2.7431$$

$$x_3^{(1)} = (1 - 1.25) \cdot 0 + 1.25 \cdot (-7 + 0.41667 + 2.7431)/3 = -1.6001.$$

The next three iterations are

$$x^{(2)} = (1.4972, 2.1880, -2.2288)^t,$$

$$x^{(3)} = (1.0494, 1.8782, -2.0141)^t,$$

$$x^{(4)} = (0.9428, 2.0007, -1.9723)^t.$$

The exact solution is $x = (1, 2, -2)^t$.

4. Error Bounds

Definition 4.1. *Suppose $\tilde{x} \in \mathbb{R}^n$ is an approximation to the solution of the linear system defined by $Ax = b$. The residual vector for $\tilde{x}$ with respect to this system is $r = b - A\tilde{x}$.*

It seems intuitively reasonable that if $\tilde{x}$ is an approximation to the solution x of $Ax = b$ and the residual vector $r = b - A\tilde{x}$ has the property that $\|r\|$ is small, then $\|x - \tilde{x}\|$ would be small as well. This is often the case, but certain systems, which occur frequently in practice, fail to have this property.

Example 9. *The linear system $Ax = b$ given by*

$$\begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix}$$

has the unique solution $x = (1, 1)^t$. Determine the residual vector for the poor approximation $\tilde{x} = (3, -0.0001)^t$.

Sol. We have

$$r = b - A\tilde{x} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -0.0001 \end{bmatrix} = \begin{bmatrix} 0.0002 \\ 0 \end{bmatrix}$$

so $\|r\|_\infty = 0.0002$. Although the norm of the residual vector is small, the approximation $\tilde{x} = (3, -0.0001)^t$ is obviously quite poor; in fact, $\|x - \tilde{x}\|_\infty = 2$.

Theorem 4.2. *Suppose that $\tilde{x}$ is an approximation to the solution of $Ax = b$, A is a nonsingular matrix, and r is the residual vector for $\tilde{x}$. Then for any natural norm,*

$$\|x - \tilde{x}\| \leq \|r\| \cdot \|A^{-1}\|$$

and if $x \neq 0$ and $b \neq 0$

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|r\|}{\|b\|}.$$

Proof. Since $r = b - A\tilde{x} = Ax - A\tilde{x}$ and A is nonsingular, we have $x - \tilde{x} = A^{-1}r$.

$$\|x - \tilde{x}\| = \|A^{-1}r\| \leq \|A^{-1}\| \cdot \|r\|.$$

Moreover, since $b = Ax$, we have $\|b\| = \|Ax\| \leq \|A\| \cdot \|x\|$. So $1/\|x\| \leq \|A\|/\|b\|$ and

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \frac{\|A\| \cdot \|A^{-1}\|}{\|b\|} \|r\|.$$

■

Condition Numbers: The inequalities in the above theorem imply that $\|A^{-1}\|$ and $\|A\| \cdot \|A^{-1}\|$ provide an indication of the connection between the residual vector and the accuracy of the approximation. In general, the relative error $\|x - \tilde{x}|/\|x\|$ is of most interest, and this error is bounded by the product of $\|A\| \cdot \|A^{-1}\|$ with the relative residual for this approximation, $\|r\|/\|b\|$. Any convenient norm can be used for this approximation; the only requirement is that it be used consistently throughout.

Definition 4.3. *The condition number of the nonsingular matrix A relative to a norm $\|\cdot\|$ is*

$$K(A) = \|A\| \cdot \|A^{-1}\|.$$

With this notation, the inequalities in above theorem become

$$\|x - \tilde{x}\| \leq K(A) \frac{\|r\|}{\|A\|}$$

and

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq K(A) \frac{\|r\|}{\|b\|}.$$

For any nonsingular matrix A and natural norm $\|\cdot\|$,

$$1 = \|I\| = \|A \cdot A^{-1}\| \leq \|A\| \cdot \|A^{-1}\| = K(A)$$

A matrix A is well-conditioned if $K(A)$ is close to 1, and is ill-conditioned when $K(A)$ is significantly greater than 1. Conditioning in this context refers to the relative security that a small residual vector implies a correspondingly accurate approximate solution. When it is very large, the solution of $Ax = b$ will be very sensitive to relatively small changes in b. Or in the the residual, a relatively small residual will quite possibly lead to a relatively large error in $\tilde{x}$ as compared with x. These comments are also valid when the changes are made to A rather than to b.

Example 10. *Determine the condition number for the matrix and comment.*

$$A = \begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix}.$$

Sol. We have $\|A\|_\infty = \max\{|1| + |2|, |1.001| + |2|\} = 3.0001$, which would not be considered large. However,

$$A^{-1} = \begin{bmatrix} -10000 & 10000 \\ 5000.5 & -5000 \end{bmatrix}, \quad \text{so} \quad \|A^{-1}\|_\infty = 20000,$$

and for the infinity norm, $K(A) = (20000)(3.0001) = 60002$. The size of the condition number for this example is very large. Thus small changes will effect the solution.

Example 11. *Suppose* $\bar{x} = \begin{bmatrix} 0.98 \\ 1.1 \end{bmatrix}$ *is an approximate solution for the linear system* $Ax = b$*, where*

$$A = \begin{bmatrix} 3.9 & 1.6 \\ 6.8 & 2.9 \end{bmatrix}, \quad b = \begin{bmatrix} 5.5 \\ 9.7 \end{bmatrix}.$$

Find a bound for the relative error $\dfrac{\|x - \bar{x}\|}{\|x\|}$.

Sol. The residual is given by

$$r = b - A\bar{x} = \begin{bmatrix} 5.5 \\ 9.7 \end{bmatrix} - \begin{bmatrix} 3.9 & 1.6 \\ 6.8 & 2.9 \end{bmatrix} \begin{bmatrix} 0.98 \\ 1.1 \end{bmatrix} = \begin{bmatrix} -0.0820 \\ -0.1540 \end{bmatrix}.$$

The bound for the relative error is (for the infinity norm)

$$\frac{\|x - \bar{x}\|}{\|x\|} \leq \frac{\|A\| \, \|A^{-1}\| \, \|r\|}{\|b\|}.$$

Also

$$\det(A) = 0.43.$$

$$\therefore A^{-1} = \frac{1}{0.43} \begin{bmatrix} 2.9 & -1.6 \\ -6.8 & 2.9 \end{bmatrix} = \begin{bmatrix} 6.7442 & -3.7209 \\ -15.8140 & 9.0698 \end{bmatrix}$$

$$\|A\| = 9.7, \ \|A^{-1}\| = 24.8837, \ \|r\| = 0.1540, \ \|b\| = 9.7.$$

$$\therefore \frac{\|x - \bar{x}\|}{\|x\|} \leq \frac{\|A\| \, \|A^{-1}\| \, \|r\|}{\|b\|} = 3.8321.$$

Example 12. *Find the condition number* $K(A)$ *of the matrix*

$$A = \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix}, \quad |c| \neq 1.$$

When does A *become ill-conditioned? What does this say about the linear system* $Ax = b$*? How is* $K(A)$ *related to* $\det(A)$*?*

Sol. We have

$$A = \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix}$$

and is well conditioned if $K(\mathrm{A})$ is near 1. $K(A)$ with respect to norm $\|\cdot\|_\infty$ is given as

$$K(A) = \|A\|_\infty \, \|A^{-1}\|_\infty.$$

Here $\det(A) = 1 - c^2$ and $adj(A) = \begin{bmatrix} 1 & -c \\ -c & 1 \end{bmatrix}$. Thus $A^{-1} = \begin{bmatrix} \frac{1}{1-c^2} & \frac{-c}{1-c^2} \\ \frac{-c}{1-c^2} & \frac{1}{1-c^2} \end{bmatrix}$.

Thus $||A||_\infty = 1 + |c|$ and $||A^{-1}||_\infty = \dfrac{1}{|1-c^2|} + \dfrac{|c|}{|1-c^2|} = \dfrac{1+|c|}{|1-c^2|}$.

Hence condition number $K(\mathrm{A}) = \dfrac{(1+|c|)^2}{|1-c^2|}$.

Thus A is ill-conditioned when $|c|$ is near 1.

When condition number is large the solution of the system $Ax = b$ is sensitive to small changes in A. If the determinant of A is small, then the condition number of A will be very large.

5. Power method for approximating eigenvalues

The eigenvalues of an $n \times n$ of matrix A are obtained by solving its characteristic equation

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

$$\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \cdots + c_0 = 0.$$

For large values of n, the polynomial equations like this one are difficult, time-consuming to solve and sensitive to rounding errors. In this section we look at an alternative method known as Power Method for approximating eigenvalues. The method is an iterative method used to determine the dominant eigenvalue - that is the eigenvalue with largest magnitude. By modifying the method it can be used to determine other eigenvalues. One useful feature of power method is that it produces not only eigenvalue but also associated eigenvector.

To apply the power method, we assume that $n \times n$ matrix A has n eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ (which we don't know) with associated eigenvectors $v^{(1)}, v^{(2)}, \cdots, v^{(n)}$. We say matrix A is diagonalizable. We write

$$Av^{(i)} = \lambda_i v^{(i)}, \quad i = 1, 2, \cdots, n.$$

We assume that these eigenvalues are ordered so that λ_1 is the dominant eigenvalue (with corresponding eigenvector $v^{(1)}$).

From linear algebra, if A is diagonalizable, then it has n linearly independent eigenvectors $v^{(1)}, v^{(2)}, \cdots, v^{(n)}$. An $n \times n$ matrix need not have n linearly independent eigenvectors. When it does not the Power method may still be successful, but it is not guaranteed to be.

As the n eigenvectors $v^{(1)}, v^{(2)}, \cdots, v^{(n)}$ are linearly independent, they must form a basis for $\mathbb{R}^n$.

We select an arbitrary nonzero starting vector $x^{(0)}$ and express it as a linear combination of basis vectors as

$$x_0 = c_1 v^{(1)} + c_2 v^{(2)} + \cdots + c_n v^{(n)}.$$

We assume that $c_1 \neq 0$. (If $c_1 = 0$, the power method may not converge, and a different $x^{(0)}$ must be used as an initial approximation.

Then we repeatedly carry out matrix-vector multiplication, using the matrix A to produce a sequence of vectors. Specifically, we have

$$\begin{aligned} x^{(1)} &= Ax^{(0)} \\ x^{(2)} &= Ax^{(1)} = A^2 x^{(0)} \\ &\vdots \\ x^{(k)} &= Ax^{(k-1)} = A^k x^{(0)}. \end{aligned}$$

In general, we have

$$x^{(k)} = A^k x^{(0)}, \quad k = 1, 2, 3, \cdots$$

Substituting the value of $x^{(0)}$, we obtain

$$\begin{aligned} x^{(k)} &= A^k x^{(0)} \\ &= c_1 A^k v^{(1)} + c_2 A^k v^{(2)} + \cdots + c_n A^k v^{(n)} \\ &= c_1 \lambda_1^k v^{(1)} + c_2 \lambda_2^k v^{(2)} + \cdots + c_n \lambda_n^k v^{(n)} \\ &= \lambda_1^k \left[c_1 v^{(1)} + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v^{(2)} + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k v^{(n)} \right] \end{aligned}$$

Now, from our original assumption that λ_1 is larger in absolute value than the other eigenvalues it follows that each of the fractions

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

Therefore each of the factors

$$\left(\frac{\lambda_2}{\lambda_1}\right)^k, \left(\frac{\lambda_3}{\lambda_1}\right)^k, \cdots, \left(\frac{\lambda_n}{\lambda_1}\right)^k$$

must approach 0 as k approaches infinity. This implies that the approximation

$$A^k x^{(0)} \approx \lambda_1^k c_1 v^{(1)}, \quad c_1 \neq 0.$$

Since $v^{(1)}$ is a dominant eigenvector, it follows that any scalar multiple of $v^{(1)}$ is also a dominant eigenvector. Thus we have shown that $A^k x_0$ approaches a multiple of the dominant eigenvector of A. The entries of $A^k x^{(0)}$ may grow with k, therefore we scale the powers of $A^k x^{(0)}$ in an appropriate manner to ensure that the limit is finite and nonzero. The scaling begins by choosing initial guess to be a unit vector $x^{(0)}$ relative to maximum norm, that is $\|x^{(0)}\|_\infty = 1$. Then we compute $y^{(1)} = Ax^{(0)}$ and next approximation can be taken as

$$x^{(1)} = \frac{y^{(1)}}{\|y^{(1)}\|_\infty}.$$

We repeat the procedure and stop by putting the following stopping criteria:

$$\frac{\|x^{(k)} - x^{(k-1)}\|_\infty}{\|x^{(k)}\|_\infty} < \varepsilon,$$

where ε is the desired accuracy.

Example 13. *Calculate four iterations of the power method with scaling to approximate a dominant eigenvector of the matrix*

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

Sol. Using $x^{(0)} = [1, 1, 1]^T$ as initial approximation, we obtain

$$y^{(1)} = Ax^{(0)} = \begin{bmatrix} 1 & 2 & 0 \\ -2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$$

and by scaling we obtain the approximation

$$x^{(1)} = 1/5 \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 0.60 \\ 0.20 \\ 1.00 \end{bmatrix}.$$

Similarly we get

$$y^{(2)} = Ax^{(1)} = \begin{bmatrix} 1.00 \\ 1.00 \\ 2.20 \end{bmatrix} = 2.20 \begin{bmatrix} 0.45 \\ 0.45 \\ 1.00 \end{bmatrix} = 2.20x^{(2)}.$$

$$y^{(3)} = Ax^{(2)} = \begin{bmatrix} 1.35 \\ 1.55 \\ 2.8 \end{bmatrix} = 2.8 \begin{bmatrix} 0.48 \\ 0.55 \\ 1.00 \end{bmatrix} = 2.8x^{(3)}.$$

$$y^{(4)} = Ax^{(3)} = 3.1 \begin{bmatrix} 0.51 \\ 0.51 \\ 1.00 \end{bmatrix}.$$

etc.
After four iterations, we observe that dominant eigenvector is

$$x = \begin{bmatrix} 0.51 \\ 0.51 \\ 1.00 \end{bmatrix}.$$

Scaling factors are approaching to dominant eigenvalue $\lambda = 3.1$.

Remark 5.1. *The power method is useful to compute the eigenvalue but it gives only dominant eigenvalue. To find other eigenvalue we use properties of matrix such as sum of all eigenvalue is equal to the trace of matrix. Also if λ is an eigenvalue of A then λ^{-1} is the eigenvalue of A^{-1}. Hence the smallest eigenvalue of A is the dominant eigenvalue of A^{-1}.*

5.1. **Inverse Power method.** The Inverse Power method is a modification of the Power method that is used to determine the eigenvalue of A that is closest to a specified number σ.
We consider $A-\sigma I$ then its eigenvalues are $\lambda_1-\sigma,\ \lambda_2-\sigma,\cdots,\lambda_n-\sigma$, where $\lambda_1,\lambda_2,\cdots,\lambda_n$ are the eigenvalues of A.
Now the eigenvalues of $(A-\sigma I)^{-1}$ are $\dfrac{1}{\lambda_1-\sigma},\ \dfrac{1}{\lambda_2-\sigma},\cdots,\dfrac{1}{\lambda_n-\sigma}$.
The eigenvalues of the original matrix A that is the closest to σ corresponds to the eigenvalue of largest magnitude of the shifted and inverted of matrix $(A-\sigma I)^{-1}$.
To find the eigenvalue closest to σ, we apply the Power method to obtain the eigenvalue μ of $(A-\sigma I)^{-1}$. Then we recover the eigenvalue λ of the original problem by $\lambda=1/\mu+\sigma$. This method is called shifted and inverted. We solve $y=(A-\sigma I)^{-1}x$ which implies $(A-\sigma I)y=x$. We need not to compute the inverse of the matrix.

Example 14. *Let λ be the smallest eigen-value in magnitude of the matrix*

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

Using four-digit rounding arithmetic and $\mathbf{x^{(0)}}=[\mathbf{1},\mathbf{0},\mathbf{0}]^{\mathbf{t}}$, perform two iterations of the inverse power method to approximate the value of λ. Use the LU factorization to solve the system of linear equations that originate during the process.

Sol. In order to get the smallest eigen-value of matrix A, we apply power method on matrix A^{-1}. Dominant eigen-value of A^{-1} will be the smallest eigen-value of matrix A. We have

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

The initial guess is $\mathbf{x^{(0)}}=[\mathbf{1},\mathbf{0},\mathbf{0}]^{\mathbf{t}}$.
First iteration:

$$\begin{aligned}A^{-1}\mathbf{x^{(0)}}&=\mathbf{y^{(1)}}\\A\mathbf{y^{(1)}}&=\mathbf{x^{(0)}}\\LU\mathbf{y^{(1)}}&=\mathbf{x^{(0)}}.\end{aligned}$$

Firstly we write LU decomposition of matrix A to solve the above system. We apply Gauss elimination on A and thus obtain

$$L=\begin{bmatrix}1&0&0\\-2&1&0\\0&-0.5&1\end{bmatrix}$$

and

$$U=\begin{bmatrix}1&-1&0\\0&2&-2\\0&0&1\end{bmatrix}.$$

Assuming $U\mathbf{y^{(1)}}=\mathbf{z^{(1)}}$, we obtain $L\mathbf{z^{(1)}}=\mathbf{x^{(0)}}$.
Solving $L\mathbf{z^{(1)}}=\mathbf{x^{(0)}}$ using forward substitution, we obtain $\mathbf{z^{(1)}}=[\mathbf{1},\mathbf{2},\mathbf{1}]^{\mathbf{t}}$,
and solving $U\mathbf{y^{(1)}}=\mathbf{z^{(1)}}$ using backward substitution, we get $\mathbf{y^{(1)}}=[\mathbf{3},\mathbf{2},\mathbf{1}]^{\mathbf{t}}$.
By scaling, we have $\mathbf{y^{(1)}}=\mathbf{3}[\mathbf{1},\mathbf{0.6667},\mathbf{0.3333}]^{\mathbf{t}}=\mu^{(1)}\mathbf{x^{(1)}}$, where $\mu^{(1)}$ is the first approximation to the dominant eigen-value of A^{-1}.
Second iteration:
Solving $L\mathbf{z^{(2)}}=\mathbf{x^{(1)}}$ using forward substitution, we obtain $\mathbf{z^{(2)}}=[\mathbf{1},\mathbf{2.6667},\mathbf{1},\mathbf{6667}]^{\mathbf{t}}$,
and solving $U\mathbf{y^{(2)}}=\mathbf{z^{(2)}}$ using backward substitution, we get $\mathbf{y^{(2)}}=[\mathbf{4},\mathbf{3},\mathbf{1.6667}]^{\mathbf{t}}$.
Therefore, we have $\mathbf{y^{(2)}}=\mathbf{4}[\mathbf{1},\mathbf{0.7500},\mathbf{0.4167}]^{\mathbf{t}}=\mu^{(2)}\mathbf{x^{(2)}}$.

So, after two iterations, the small eigen-value of A is

$$\begin{aligned}\lambda \approx & \frac{1}{\mu^{(2)}} \\ = & 0.2500\end{aligned}$$

and the corresponding eigen-vector is $\mathbf{x} \approx [\mathbf{1}, \mathbf{0.7500}, \mathbf{0.4167}]^{\mathbf{t}}$.

Example 15. *Apply the inverse power method with* $x^{(0)} = [1, 1, 1]^T$ *to the matrix*

$$\begin{bmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

to find eigenvalue near to $19/3$.

Sol. For the inverse power method, we consider

$$A - \frac{19}{3}I = \begin{bmatrix} \frac{-31}{3} & 14 & 0 \\ -5 & \frac{20}{3} & 0 \\ -1 & 0 & -\frac{13}{3} \end{bmatrix}$$

Starting with $x^{(0)} = [1, 1, 1]^T$, $(A - \sigma I)^{-1}x^{(0)} = y^{(1)}$ gives $(A - \sigma I)y^{(1)} = x^{(0)}$. This gives

$$\begin{bmatrix} \frac{-31}{3} & 14 & 0 \\ -5 & \frac{20}{3} & 0 \\ -1 & 0 & -\frac{13}{3} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Solving above system by Gauss elimination (LU decomposition), we get $a = -6.6$, $b = -4.8$, and $c = 1.2923$.

Therefore $y^{(1)} = (-6.6, -4.8, 1.2923)^T$. We normalize it by taking 6.6 as scale factor and $x^{(1)} = \frac{1}{-6.6}y^{(1)} = (1, 0.7272, -0.1958)^T$.

Therefore first approximation of the eigenvalue of A near $19/3$ is $-\frac{1}{6.6} + \frac{19}{3} = 6.1818$.

Repeating the above procedure we can obtain the eigenvalue (and which is 6).

Important Remark: Although the power method worked well in these examples, we must say something about cases in which the power method may fail. There are basically three such cases:

1. Using the power method when A is not diagonalizable. Recall that A has n linearly Independent eigenvector if and only if A is diagonalizable. Of course, it is not easy to tell by just looking at A whether it is diagonalizable.
2. Using the power method when A does not have a dominant eigenvalue or when the dominant eigenvalue is such that $|\lambda_1| = |\lambda_2|$.
3. If the entries of A contains significant error. Powers A^k will have significant roundoff error in their entires.

Exercises

(1) Find l_∞ and l_2 norms of the vectors.
   (a) $x = (3, -4, 0, \frac{3}{2})^t$.
   (b) $x = (\sin k, \cos k, 2^k)^t$ for a fixed positive integer k.

(2) Find the l_∞ norm of the matrix: $\begin{bmatrix} 4 & -1 & 7 \\ -1 & 4 & 0 \\ -7 & 0 & 4 \end{bmatrix}$.

(3) The following linear system $Ax = b$ have x as the actual solution and $\bar{x}$ as an approximate solution. Compute $\|x - \bar{x}\|_\infty$ and $\|A\bar{x} - b\|_\infty$. Also compute $\|A\|_\infty$.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 + 3x_2 + 4x_3 &= -1 \\ 3x_1 + 4x_2 + 6x_3 &= 2, \\ x &= (0, -7, 5)^t \\ \bar{x} &= (-0.2, -7.5, 5.4)^t. \end{aligned}$$

(4) Find the first two iterations of Jacobi and Gauss-Seidel using $x^{(0)} = 0$:

$$\begin{aligned} 4.63x_1 - 1.21x_2 + 3.22x_3 &= 2.22 \\ -3.07x_1 + 5.48x_2 + 2.11x_3 &= -3.17 \\ 1.26x_1 + 3.11x_2 + 4.57x_3 &= 5.11. \end{aligned}$$

(5) The linear system

$$\begin{aligned} x_1 - x_3 &= 0.2 \\ -\frac{1}{2}x_1 + x_2 - \frac{1}{4}x_3 &= -1.425 \\ x_1 - \frac{1}{2}x_2 + x_3 &= 2 \end{aligned}$$

has the solution $(0.9, -0.8, 0.7)^T$.
   (a) Is the coefficient matrix strictly diagonally dominant?
   (b) Compute the spectral radius of the Gauss-Seidel iteration matrix.
   (c) Perform four iterations of the Gauss-Seidel iterative method to approximate the solution.
   (d) What happens in part (c) when the first equation in the system is changed to $x_1 - 2x_3 = 0.2$?

(6) Comment whether you can apply Gauss-Seidel method for the following system of equations. If so, then perform two iterations by taking initial guess as $x^{(0)} = [0, 0, 0]^t$.

$$\begin{aligned} 12x_1 + 3x_2 - 5x_3 &= 1 \\ x_1 + 5x_2 + 3x_3 &= 28 \\ 3x_1 + 7x_2 + 13x_3 &= 76. \end{aligned}$$

(7) Find the first two iterations of the SOR method with $\omega = 1.1$ for the following linear systems, using $x^{(0)} = [0, 0, 0[^t$

$$\begin{aligned} 4x_1 + x_2 - x_3 &= 5 \\ -x_1 + 3x_2 + x_3 &= -4 \\ 2x_1 + 2x_2 + 5x_3 &= 1. \end{aligned}$$

(8) Compute the condition numbers of the following matrices relative to $\|.\|_\infty$.
   (a) $\begin{bmatrix} 0.03 & 58.9 \\ 5.31 & -6.10 \end{bmatrix}$
   (b) $\begin{bmatrix} 0.04 & 0.01 & -0.01 \\ 0.2 & 0.5 & -0.2 \\ 1 & 2 & 4 \end{bmatrix}$.

(9) The linear system $Ax = b$ given by

$$\begin{bmatrix} 1 & 2 \\ 1.0001 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.0001 \end{bmatrix}$$

has solution $(1, 1)^t$. Use four-digit rounding arithmetic to find the solution of the perturbed system

$$\begin{bmatrix} 1 & 2 \\ 1.000011 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3.00001 \\ 3.00003 \end{bmatrix}$$

Is matrix A ill-conditioned?

(10) Determine the dominant eigenvalue and the corresponding eigenvector of the following matrix using the power method with $x^{(0)} = [1, -1, 2]^t$.

$$\begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}.$$

Use stopping criteria with absolute error less than 0.001 for eigenvalue.

(11) Use the inverse power method to approximate the smallest eigenvalue of the matrix until a tolerance of 10^{-2} is achieved with $x^{(0)} = [1, 0, 0]^t$.

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

(12) Find the eigenvalue of matrix nearest to 3

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

using inverse power method.

Appendix A. Algorithms

Algorithm (Gauss-Seidel):

(1) Input matrix $A = [a_{ij}], b, XO = x^{(0)}$, tolerance TOL, maximum number of iterations
(2) Set $k = 1$
(3) while $(k \leq N)$ do step 4-7
(4) For $i = 1, 2, \cdots, n$

$$x_i = \frac{1}{a_{ii}} \left[-\sum_{j=1}^{i-1} (a_{ij} x_j) - \sum_{j=i+1}^{n} (a_{ij} XO_j) + b_i) \right]$$

(5) If $||x - XO|| < TOL$, then OUTPUT $(x_1, x_2, \cdots, x_n)$
STOP
(6) $k = k + 1$
(7) For $i = 1, 2, \cdots, n$
Set $XO_i = x_i$
(8) OUTPUT $(x_1, x_2, \cdots, x_n)$
STOP.

Algorithm (Power Method):

(1) Start
(2) Define matrix A and initial guess x
(3) Calculate $y = Ax$
(4) Find the largest element in magnitude of matrix y and assign it to K.
(5) Calculate fresh value $x = (1/K) * y$
(6) If $[K(n) - K(n-1)] >$ error, goto step 3.
(7) Stop

Bibliography

[Burden] Richard L. Burden, J. Douglas Faires and Annette Burden, "Numerical Analysis," Cengage Learning, 10th edition, 2015.

[Atkinson] K. Atkinson and W. Han, "Elementary Numerical Analysis," John Willey and Sons, 3rd edition, 2004.

[Cheney] E. Ward Cheney and David R. Kincaid, "Numerical Mathematics and Computing", Cengage Learning, 7th edition, 2012.