7.3 The Jacobi and Gauss-Seidel Iterative Methods
The Jacobi Method

Two assumptions made on Jacobi Method:

1. The system given by
   \[ a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \]
   \[ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \]
   \[ \vdots \]
   \[ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \]

   Has a unique solution.

2. The coefficient matrix \( A \) has no zeros on its main diagonal, namely, \( a_{11}, a_{22}, \ldots, a_{nn} \) are nonzeros.
Main idea of Jacobi
To begin, solve the 1\textsuperscript{st} equation for $x_1$, the 2\textsuperscript{nd} equation for $x_2$ and so on to obtain the rewritten equations:

\[ x_1 = \frac{1}{a_{11}} (b_1 - a_{12}x_2 - a_{13}x_3 - \cdots - a_{1n}x_n) \]
\[ x_2 = \frac{1}{a_{22}} (b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2n}x_n) \]
\[ \vdots \]
\[ x_n = \frac{1}{a_{nn}} (b_n - a_{n1}x_1 - a_{n2}x_2 - \cdots - a_{nn-1}x_{n-1}) \]

Then make an initial guess of the solution \( x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \ldots, x_n^{(0)}) \). Substitute these values into the right hand side of the rewritten equations to obtain the \textit{first approximation}, \( (x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \ldots, x_n^{(1)}) \).

This accomplishes one \textit{iteration}. 
In the same way, the second approximation \( (x_1^{(2)}, x_2^{(2)}, x_3^{(2)}, \ldots, x_n^{(2)}) \) is computed by substituting the first approximation’s value \( (x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \ldots, x_n^{(1)}) \) into the right hand side of the rewritten equations. By repeated iterations, we form a sequence of approximations \( x^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, \ldots, x_n^{(k)})^t, \quad k = 1, 2, 3, \ldots \)

**The Jacobi Method.** For each \( k \geq 1 \), generate the components \( x_i^{(k)} \) of \( x^{(k)} \) from \( x^{(k-1)} \) by

\[
x_i^{(k)} = \frac{1}{a_{ii}} \left[ \sum_{\substack{j=1, \ j \neq i}}^{n} (-a_{ij} x_j^{(k-1)}) + b_i \right], \quad \text{for } i = 1, 2, \ldots, n
\]
Example. Apply the Jacobi method to solve

\[
\begin{align*}
5x_1 - 2x_2 + 3x_3 &= -1 \\
-3x_1 + 9x_2 + x_3 &= 2 \\
2x_1 - x_2 - 7x_3 &= 3
\end{align*}
\]

Continue iterations until two successive approximations are identical when rounded to three significant digits.

Solution

<table>
<thead>
<tr>
<th>$n$</th>
<th>$k = 0$</th>
<th>$k = 1$</th>
<th>$k = 2$</th>
<th>$k = 3$</th>
<th>$k = 4$</th>
<th>$k = 5$</th>
<th>$k = 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1^{(k)}$</td>
<td>0.000</td>
<td>-0.200</td>
<td>0.146</td>
<td>0.192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_2^{(k)}$</td>
<td>0.000</td>
<td>0.222</td>
<td>0.203</td>
<td>0.328</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_3^{(k)}$</td>
<td>0.000</td>
<td>-0.429</td>
<td>-0.517</td>
<td>-0.416</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When to stop: 1. \( \frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|} < \varepsilon \); or \( \|x^{(k)} - x^{(k-1)}\| < \varepsilon \). Here \( \varepsilon \) is a given small number. Another stopping criterion: \( \frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|} \)

Definition 7.1 A vector norm on \( \mathbb{R}^n \) is a function, \( \| \cdot \| \), from \( \mathbb{R}^n \) to \( \mathbb{R} \) with the properties:

(i) \( \|x\| \geq 0 \) for all \( x \in \mathbb{R}^n \)
(ii) \( \|x\| = 0 \) if and only if \( x = 0 \)
(iii) \( \|\alpha x\| = |\alpha|\|x\| \) for all \( \alpha \in \mathbb{R} \) and \( x \in \mathbb{R}^n \)
(iv) \( \|x + y\| \leq \|x\| + \|y\| \) for all \( x, y \in \mathbb{R}^n \)

Definition 7.2 The Euclidean norm \( l_2 \) and the infinity norm \( l_{\infty} \) for the vector \( x = [x_1, x_2, \ldots, x_n]^t \) are defined by

\[
\|x\|_2 = \left\{ \sum_{i=1}^{n} x_i^2 \right\}^{\frac{1}{2}}
\]

and
\[ \| \mathbf{x} \|_\infty = \max_{1 \leq i \leq n} |x_i| \]

**Example.** Determine the \( l_2 \) and \( l_\infty \) norms of the vector \( \mathbf{x} = (-1, 1, -2)^t \).

**Solution:**

\[ \| \mathbf{x} \|_2 = \sqrt{(-1)^2 + (1)^2 + (-2)^2} = \sqrt{6}. \]

\[ \| \mathbf{x} \|_\infty = \max\{|-1|, |1|, |-2|\} = 2. \]
The Jacobi Method in Matrix Form

Consider to solve an $n \times n$ size system of linear equations $Ax = b$ with

$$A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \text{for} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$ 

We split $A$ into
\[ A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} - \begin{bmatrix} 0 & \cdots & 0 & 0 \\ -a_{21} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ -a_{n1} & \cdots & -a_{n,n-1} & 0 \end{bmatrix} - \begin{bmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & -a_{n-1,n} \\ 0 & 0 & \cdots & 0 \end{bmatrix} = D - L - U \]

Where \( D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \), \( L = \begin{bmatrix} -a_{21} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ -a_{n1} & \cdots & -a_{n,n-1} & 0 \end{bmatrix} \), and \( U \) is a similar matrix with \( a_{nn} \) in the bottom right corner.
\[ U = \begin{bmatrix}
0 & -a_{12} & \ldots & -a_{1n} \\
0 & 0 & \ddots & \vdots \\
\vdots & \vdots & \ddots & -a_{n-1,n} \\
0 & 0 & \ldots & 0
\end{bmatrix} \]

\[ Ax = b \text{ is transformed into } (D - L - U)x = b. \]

\[ Dx = (L + U)x + b \]

Assume \( D^{-1} \) exists and \( D^{-1} = \begin{bmatrix}
\frac{1}{a_{11}} & 0 & \ldots & 0 \\
0 & \frac{1}{a_{22}} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \frac{1}{a_{nn}}
\end{bmatrix} \]

Then

\[ x = D^{-1}(L + U)x + D^{-1}b \]

The matrix form of Jacobi iterative method is

\[ x^{(k)} = D^{-1}(L + U)x^{(k-1)} + D^{-1}b \quad k = 1, 2, 3, \ldots \]
Define $T_j = D^{-1}(L + U)$ and $c = D^{-1}b$, Jacobi iteration method can also be written as

$$x^{(k)} = T_j x^{(k-1)} + c \quad k = 1, 2, 3, \ldots$$

**The Gauss-Seidel Method**

**Main idea of Gauss-Seidel**
With the Jacobi method, only the values of $x^{(k)}_i$ obtained in the $k$th iteration are used to compute $x^{(k+1)}_i$. With the Gauss-Seidel method, we use the new values $x^{(k+1)}_i$ as soon as they are known. For example, once we have computed $x^{(k+1)}_1$ from the first equation, its value is then used in the second equation to obtain the new $x^{(k+1)}_2$, and so on.

Example. Use the Gauss-Seidel method to solve
\[ 5x_1 - 2x_2 + 3x_3 = -1 \]
\[ -3x_1 + 9x_2 + x_3 = 2 \]
\[ 2x_1 - x_2 - 7x_3 = 3 \]

Choose the initial guess \( x_1 = 0, x_2 = 0, x_3 = 0 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( k = 0 )</th>
<th>( k = 1 )</th>
<th>( k = 2 )</th>
<th>( k = 3 )</th>
<th>( k = 4 )</th>
<th>( k = 5 )</th>
<th>( k = 6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1^{(k)} )</td>
<td>0.000</td>
<td>-0.200</td>
<td>0.167</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( x_2^{(k)} )</td>
<td>0.000</td>
<td>0.156</td>
<td>0.334</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( x_2^{(k)} )</td>
<td>0.000</td>
<td>-0.508</td>
<td>-0.429</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**The Gauss-Seidel Method.** For each \( k \geq 1 \), generate the components \( x_i^{(k)} \) of \( x^{(k)} \) from \( x^{(k-1)} \) by
\[
x_i^{(k)} = \frac{1}{a_{ii}} \left[ \sum_{j=1}^{i-1} (a_{ij} x_j^{(k)}) - \sum_{j=i+1}^{n} (a_{ij} x_j^{(k-1)}) + b_i \right], \quad \text{for } i = 1, 2, \ldots, n
\]

Namely,
\[
\begin{align*}
    a_{11} x_1^{(k)} &= -a_{12} x_2^{(k-1)} - \cdots - a_{1n} x_n^{(k-1)} + b_1 \\
    a_{22} x_2^{(k)} &= -a_{21} x_1^{(k)} - a_{23} x_3^{(k-1)} - \cdots - a_{2n} x_n^{(k-1)} + b_2 \\
    a_{33} x_3^{(k)} &= -a_{31} x_1^{(k)} - a_{32} x_2^{(k)} - a_{34} x_4^{(k-1)} - \cdots - a_{3n} x_n^{(k-1)} + b_3 \\
    &\vdots \\
    a_{nn} x_n^{(k)} &= -a_{n1} x_1^{(k)} - a_{n2} x_2^{(k)} - \cdots - a_{n,n-1} x_{n-1}^{(k)} + b_n
\end{align*}
\]

Matrix form of Gauss-Seidel method.
\[(D - L)x^{(k)} = Ux^{(k-1)} + b\]

\[x^{(k)} = (D - L)^{-1}Ux^{(k-1)} + (D - L)^{-1}b\]

Define \(T_g = (D - L)^{-1}U\) and \(c_g = (D - L)^{-1}b\), Gauss-Seidel method can be written as

\[x^{(k)} = T_g x^{(k-1)} + c_g \quad k = 1,2,3, ...\]

Convergence theorems of the iteration methods
Let the iteration method be written as
\[ x^{(k)} = T x^{(k-1)} + c \]
for each \( k = 1, 2, 3, \ldots \)

**Definition 7.14** The **spectral radius** \( \rho(A) \) of a matrix \( A \) is defined by
\[ \rho(A) = \max |\lambda|, \quad \text{where} \ \lambda \ \text{is an eigenvalue of} \ A. \]

Remark: For complex \( \lambda = a + bj \), we define \(|\lambda| = \sqrt{a^2 + b^2}\).

**Lemma 7.18** If the spectral radius satisfies \( \rho(T) < 1 \), then \( (I - T)^{-1} \) exists, and
\[
(I - T)^{-1} = I + T + T^2 + \cdots = \sum_{j=0}^{\infty} T^j
\]

**Theorem 7.19** For any \( x^{(0)} \in R^n \), the sequence \( \{x^{(k)}\}_{k=0}^{\infty} \) defined by
\[ x^{(k)} = Tx^{(k-1)} + c \quad \text{for each } k \geq 1 \]

converges to the unique solution of \( x = Tx + c \) if and only if \( \rho(T) < 1 \).

**Proof** (only show \( \rho(T) < 1 \) is sufficient condition)

\[
x^{(k)} = Tx^{(k-1)} + c = T(Tx^{(k-2)} + c) + c = \cdots = T^k x^{(0)} + (T^{k-1} + \cdots + T + I)c
\]

Since \( \rho(T) < 1 \), \( \lim_{k \to \infty} T^k x^{(0)} = 0 \)

\[
\lim_{k \to \infty} x^{(k)} = 0 + \lim_{k \to \infty} \left( \sum_{j=0}^{k-1} T^j \right) c = (I - T)^{-1} c
\]
Definition 7.8 A matrix norm $\| \cdot \|$ on $n \times n$ matrices is a real-valued function satisfying

(i) $\| A \| \geq 0$

(ii) $\| A \| = 0$ if and only if $A = 0$

(iii) $\| \alpha A \| = |\alpha| \| A \|$

(iv) $\| A + B \| \leq \| A \| + \| B \|$

(v) $\| AB \| \leq \| A \| \| B \|$

Theorem 7.9. If $\| \cdot \|$ is a vector norm, the induced (or natural) matrix norm is given by

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

Example. $\| A \|_\infty = \max_{\| x \|_\infty = 1} \| Ax \|_\infty$, the $l_\infty$ induced norm.

$\| A \|_2 = \max_{\| x \|_2 = 1} \| Ax \|_2$, the $l_2$ induced norm.
Theorem 7.11. 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. Determine \( \|A\|_\infty \) for the matrix \( A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{bmatrix} \)

Corollary 7.20  If \( \|T\| < 1 \) for any natural matrix norm and \( c \) is a given vector, then the sequence \( \{x^{(k)}\}_{k=0}^{\infty} \) defined by

\[ x^{(k)} = Tx^{(k-1)} + c \]

converges, for any \( x^{(0)} \in R^n \), to a vector \( x \in R^n \), with \( x = Tx + c \), and the following error bound hold:

(i) \( \|x - x^{(k)}\| \leq \|T\|^k \|x^{(0)} - x\| \)
Theorem 7.21 If $A$ is strictly diagonally dominant, then for any choice of $x^{(0)}$, both the Jacobi and Gauss-Seidel methods give sequences $\{x^{(k)}\}_{k=0}^{\infty}$ that converges to the unique solution of $Ax = b$.

Rate of Convergence

Corollary 7.20 (i) implies $\|x - x^{(k)}\| \approx \rho(T)^k \|x^{(0)} - x\|$

Theorem 7.22 (Stein-Rosenberg) If $a_{ij} \leq 0$, for each $i \neq j$ and $a_{ii} \geq 0$, for each $i = 1, 2, \ldots, n$, then one and only one of following statements holds:

(i) $0 \leq \rho(T_g) < \rho(T_j) < 1$;
(ii) $1 < \rho(T_j) < \rho(T_g)$;
(iii) $\rho(T_j) = \rho(T_g) = 0$;
(iv) $\rho(T_j) = \rho(T_g) = 1$. 