Outline

1. The Gauss-Seidel Method
Outline

1. The Gauss-Seidel Method
2. The Gauss-Seidel Algorithm
Outline

1. The Gauss-Seidel Method
2. The Gauss-Seidel Algorithm
3. Convergence Results for General Iteration Methods
Outline

1. The Gauss-Seidel Method
2. The Gauss-Seidel Algorithm
3. Convergence Results for General Iteration Methods
4. Application to the Jacobi & Gauss-Seidel Methods
Outline

1. The Gauss-Seidel Method
2. The Gauss-Seidel Algorithm
3. Convergence Results for General Iteration Methods
4. Application to the Jacobi & Gauss-Seidel Methods
Looking at the Jacobi Method

A possible improvement to the Jacobi Algorithm can be seen by re-considering

\[
x_i^{(k)} = \frac{1}{a_{ii}} \left[ \sum_{j=1}^{n} \left( -a_{ij}x_j^{(k-1)} \right) + b_i \right], \quad \text{for } i = 1, 2, \ldots, n
\]
The Gauss-Seidel Method

Looking at the Jacobi Method

- A possible improvement to the Jacobi Algorithm can be seen by re-considering

\[
x_i^{(k)} = \frac{1}{a_{ii}} \left[ \sum_{\substack{j=1 \atop j \neq i}}^{n} (-a_{ij}x_j^{(k-1)}) + b_i \right], \quad \text{for } i = 1, 2, \ldots, n
\]

- The components of \( x^{(k-1)} \) are used to compute all the components \( x_i^{(k)} \) of \( x^{(k)} \).
The Gauss-Seidel Method

Looking at the Jacobi Method

A possible improvement to the Jacobi Algorithm can be seen by re-considering

\[ x_i^{(k)} = \frac{1}{a_{ii}} \left[ \sum_{j=1}^{n} (-a_{ij}x_j^{(k-1)}) + b_i \right], \quad \text{for } i = 1, 2, \ldots, n \]

The components of \( x^{(k-1)} \) are used to compute all the components \( x_i^{(k)} \) of \( x^{(k)} \).

But, for \( i > 1 \), the components \( x_1^{(k)}, \ldots, x_{i-1}^{(k)} \) of \( x^{(k)} \) have already been computed and are expected to be better approximations to the actual solutions \( x_1, \ldots, x_{i-1} \) than are \( x_1^{(k-1)}, \ldots, x_{i-1}^{(k-1)} \).
The Gauss-Seidel Method

Instead of using

\[ x_i^{(k)} = \frac{1}{a_{ii}} \left[ \sum_{j=1}^{n} \left( -a_{ij} x_j^{(k-1)} \right) + b_i \right] , \quad \text{for } i = 1, 2, \ldots, n \]

it seems reasonable, then, to compute \( x_i^{(k)} \) using these most recently calculated values.
The Gauss-Seidel Method

Instead of using

\[ x_i^{(k)} = \frac{1}{a_{ii}} \left[ \sum_{j=1}^{n} \left( -a_{ij}x_j^{(k-1)} \right) + b_i \right], \quad \text{for } i = 1, 2, \ldots, n \]

it seems reasonable, then, to compute \( x_i^{(k)} \) using these most recently calculated values.

The Gauss-Seidel Iterative Technique

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

for each \( i = 1, 2, \ldots, n \).
The Gauss-Seidel Method

Example

Use the Gauss-Seidel iterative technique to find approximate solutions to

\[
\begin{align*}
10x_1 - x_2 + 2x_3 &= 6 \\
-x_1 + 11x_2 - x_3 + 3x_4 &= 25 \\
2x_1 - x_2 + 10x_3 - x_4 &= -11' \\
3x_2 - x_3 + 8x_4 &= 15
\end{align*}
\]

Starting with \(x = (0, 0, 0, 0)\) and iterating until \(\|x(k) - x(k-1)\|_\infty < 10^{-3}\). Note: The solution \(x = (1, 2, -1, 1)\) was approximated by Jacobi's method in an earlier example.
The Gauss-Seidel Method

Example

Use the Gauss-Seidel iterative technique to find approximate solutions to

\[
\begin{align*}
10x_1 - x_2 + 2x_3 &= 6 \\
-x_1 + 11x_2 - x_3 + 3x_4 &= 25 \\
2x_1 - x_2 + 10x_3 - x_4 &= -11' \\
3x_2 - x_3 + 8x_4 &= 15
\end{align*}
\]

starting with \( \mathbf{x} = (0, 0, 0, 0)^t \)
The Gauss-Seidel Method

Example

Use the Gauss-Seidel iterative technique to find approximate solutions to

\[
\begin{align*}
10x_1 - x_2 + 2x_3 &= 6 \\
-x_1 + 11x_2 - x_3 + 3x_4 &= 25 \\
2x_1 - x_2 + 10x_3 - x_4 &= -11' \\
3x_2 - x_3 + 8x_4 &= 15
\end{align*}
\]

starting with \( \mathbf{x} = (0, 0, 0, 0)^t \) and iterating until

\[
\frac{\| \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} \|_\infty}{\| \mathbf{x}^{(k)} \|_\infty} < 10^{-3}
\]

Note: The solution \( \mathbf{x} = (1, 2, -1, 1)^t \) was approximated by Jacobi's method in an earlier example.
The Gauss-Seidel Method

Example

Use the Gauss-Seidel iterative technique to find approximate solutions to

\[
\begin{align*}
10x_1 - x_2 + 2x_3 &= 6 \\
-x_1 + 11x_2 - x_3 + 3x_4 &= 25 \\
2x_1 - x_2 + 10x_3 - x_4 &= -11' \\
3x_2 - x_3 + 8x_4 &= 15
\end{align*}
\]

starting with \( \mathbf{x} = (0, 0, 0, 0)^t \) and iterating until

\[
\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty}{\|\mathbf{x}^{(k)}\|_\infty} < 10^{-3}
\]

Note: The solution \( \mathbf{x} = (1, 2, -1, 1)^t \) was approximated by Jacobi’s method in an earlier example.
The Gauss-Seidel Method

Solution (1/3)

For the Gauss-Seidel method we write the system, for each \( k = 1, 2, \ldots \) as

\[
\begin{align*}
    x_1^{(k)} &= \frac{1}{10} x_2^{(k-1)} - \frac{1}{5} x_3^{(k-1)} + \frac{3}{5} \\
    x_2^{(k)} &= \frac{1}{11} x_1^{(k)} + \frac{1}{11} x_3^{(k-1)} - \frac{3}{11} x_4^{(k-1)} + \frac{25}{11} \\
    x_3^{(k)} &= -\frac{1}{5} x_1^{(k)} + \frac{1}{10} x_2^{(k)} + \frac{1}{10} x_4^{(k-1)} - \frac{11}{10} \\
    x_4^{(k)} &= -\frac{3}{8} x_2^{(k)} + \frac{1}{8} x_3^{(k)} + \frac{15}{8}
\end{align*}
\]
The Gauss-Seidel Method

### Solution (2/3)

When \( \mathbf{x}^{(0)} = (0, 0, 0, 0)^t \), we have 
\[ \mathbf{x}^{(1)} = (0.6000, 2.3272, -0.9873, 0.8789)^t. \]

<table>
<thead>
<tr>
<th>( k )</th>
<th>( x^{(k)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0, 0, 0, 0)</td>
</tr>
<tr>
<td>1</td>
<td>(0.6000, 2.3272, -0.9873, 0.8789)</td>
</tr>
<tr>
<td>2</td>
<td>(0.6000, 2.3272, -0.9873, 0.8789)</td>
</tr>
<tr>
<td>3</td>
<td>(0.6000, 2.3272, -0.9873, 0.8789)</td>
</tr>
<tr>
<td>4</td>
<td>(0.6000, 2.3272, -0.9873, 0.8789)</td>
</tr>
</tbody>
</table>
Solution (2/3)

When $\mathbf{x}^{(0)} = (0, 0, 0, 0)^t$, we have $\mathbf{x}^{(1)} = (0.6000, 2.3272, -0.9873, 0.8789)^t$. Subsequent iterations give the values in the following table:

<table>
<thead>
<tr>
<th>$k$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1^{(k)}$</td>
<td>0.0000</td>
<td>0.6000</td>
<td>1.030</td>
<td>1.0065</td>
<td>1.0009</td>
<td>1.0001</td>
</tr>
<tr>
<td>$x_2^{(k)}$</td>
<td>0.0000</td>
<td>2.3272</td>
<td>2.037</td>
<td>2.0036</td>
<td>2.0003</td>
<td>2.0000</td>
</tr>
<tr>
<td>$x_3^{(k)}$</td>
<td>0.0000</td>
<td>-0.9873</td>
<td>-1.014</td>
<td>-1.0025</td>
<td>-1.0003</td>
<td>-1.0000</td>
</tr>
<tr>
<td>$x_4^{(k)}$</td>
<td>0.0000</td>
<td>0.8789</td>
<td>0.984</td>
<td>0.9983</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
The Gauss-Seidel Method

Solution (3/3)

Because

\[
\frac{\|\mathbf{x}^{(5)} - \mathbf{x}^{(4)}\|_\infty}{\|\mathbf{x}^{(5)}\|_\infty} = \frac{0.0008}{2.000} = 4 \times 10^{-4}
\]

\(\mathbf{x}^{(5)}\) is accepted as a reasonable approximation to the solution.
The Gauss-Seidel Method

Solution (3/3)

Because

$$\frac{\|x^{(5)} - x^{(4)}\|_{\infty}}{\|x^{(5)}\|_{\infty}} = \frac{0.0008}{2.000} = 4 \times 10^{-4}$$

$x^{(5)}$ is accepted as a reasonable approximation to the solution.

Note that, in an earlier example, Jacobi’s method required twice as many iterations for the same accuracy.
The Gauss-Seidel Method: Matrix Form

Re-Writing the Equations

To write the Gauss-Seidel method in matrix form,
The Gauss-Seidel Method: Matrix Form

Re-Writing the Equations

To write the Gauss-Seidel method in matrix form, multiply both sides of

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

by \( a_{ii} \) and collect all \( k \)th iterate terms,
The Gauss-Seidel Method: Matrix Form

Re-Writing the Equations

To write the Gauss-Seidel method in matrix form, multiply both sides of

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

by \( a_{ii} \) and collect all \( k \)th iterate terms, to give

\[ a_{i1} x_1^{(k)} + a_{i2} x_2^{(k)} + \cdots + a_{ii} x_i^{(k)} = -a_{i,i+1} x_{i+1}^{(k-1)} - \cdots - a_{in} x_n^{(k-1)} + b_i \]

for each \( i = 1, 2, \ldots, n \).
The Gauss-Seidel Method: Matrix Form

Re-Writing the Equations (Cont’d)

Writing all \( n \) equations gives

\[
\begin{align*}
    a_{11}x_1^{(k)} & = -a_{12}x_2^{(k-1)} - a_{13}x_3^{(k-1)} - \cdots - a_{1n}x_n^{(k-1)} + b_1 \\
    a_{21}x_1^{(k)} + a_{22}x_2^{(k)} & = -a_{23}x_3^{(k-1)} - \cdots - a_{2n}x_n^{(k-1)} + b_2 \\
    \vdots & \\
    a_{n1}x_1^{(k)} + a_{n2}x_2^{(k)} + \cdots + a_{nn}x_n^{(k)} & = b_n
\end{align*}
\]
The Gauss-Seidel Method: Matrix Form

Re-Writing the Equations (Cont’d)

Writing all $n$ equations gives

\[
\begin{align*}
a_{11}x_1^{(k)} & = -a_{12}x_2^{(k-1)} - a_{13}x_3^{(k-1)} - \cdots - a_{1n}x_n^{(k-1)} + b_1 \\
a_{21}x_1^{(k)} + a_{22}x_2^{(k)} & = -a_{23}x_3^{(k-1)} - \cdots - a_{2n}x_n^{(k-1)} + b_2 \\
& \vdots \\
a_{n1}x_1^{(k)} + a_{n2}x_2^{(k)} + \cdots + a_{nn}x_n^{(k)} & = b_n
\end{align*}
\]

With the definitions of $D$, $L$, and $U$ given previously, we have the Gauss-Seidel method represented by

\[
(D - L)x^{(k)} = Ux^{(k-1)} + b
\]
The Gauss-Seidel Method: Matrix Form

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

Re-Writing the Equations (Cont’d)

Solving for \(x^{(k)}\) finally gives

\[x^{(k)} = (D - L)^{-1}Ux^{(k-1)} + (D - L)^{-1}b, \quad \text{for each } k = 1, 2, \ldots\]
The Gauss-Seidel Method: Matrix Form

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

Re-Writing the Equations (Cont’d)

Solving for \(x^{(k)}\) finally gives

\[x^{(k)} = (D - L)^{-1} Ux^{(k-1)} + (D - L)^{-1} b, \quad \text{for each } k = 1, 2, \ldots\]

Letting \(T_g = (D - L)^{-1} U\) and \(c_g = (D - L)^{-1} b\), gives the Gauss-Seidel technique the form

\[x^{(k)} = T_g x^{(k-1)} + c_g\]
The Gauss-Seidel Method: Matrix Form

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

Re-Writing the Equations (Cont’d)

Solving for \(x^{(k)}\) finally gives

\[x^{(k)} = (D - L)^{-1}Ux^{(k-1)} + (D - L)^{-1}b, \quad \text{for each } k = 1, 2, \ldots\]

Letting \(T_g = (D - L)^{-1}U\) and \(c_g = (D - L)^{-1}b\), gives the Gauss-Seidel technique the form

\[x^{(k)} = T_g x^{(k-1)} + c_g\]

For the lower-triangular matrix \(D - L\) to be nonsingular, it is necessary and sufficient that \(a_{ii} \neq 0, \text{ for each } i = 1, 2, \ldots, n.\)
Outline

1. The Gauss-Seidel Method
2. The Gauss-Seidel Algorithm
3. Convergence Results for General Iteration Methods
4. Application to the Jacobi & Gauss-Seidel Methods
To solve $Ax = b$ given an initial approximation $x^{(0)}$: 
Gauss-Seidel Iterative Algorithm (1/2)

To solve $Ax = b$ given an initial approximation $x^{(0)}$:

**INPUT**
- the number of equations and unknowns $n$;
- the entries $a_{ij}$, $1 \leq i, j \leq n$ of the matrix $A$;
- the entries $b_i$, $1 \leq i \leq n$ of $b$;
- the entries $XO_i$, $1 \leq i \leq n$ of $XO = x^{(0)}$;
- tolerance $TOL$;
- maximum number of iterations $N$. 

**OUTPUT**
- the approximate solution $x_1, \ldots, x_n$ or a message that the number of iterations was exceeded.
Gauss-Seidel Iterative Algorithm (1/2)

To solve \( Ax = b \) given an initial approximation \( x^{(0)} \):

**INPUT**
- the number of equations and unknowns \( n \);
- the entries \( a_{ij}, 1 \leq i, j \leq n \) of the matrix \( A \);
- the entries \( b_i, 1 \leq i \leq n \) of \( b \);
- the entries \( XO_i, 1 \leq i \leq n \) of \( XO = x^{(0)} \);
- tolerance \( TOL \);
- maximum number of iterations \( N \).

**OUTPUT**
- the approximate solution \( x_1, \ldots, x_n \) or a message that the number of iterations was exceeded.
Gauss-Seidel Iterative Algorithm (2/2)

Step 1 Set $k = 1$
Step 2 While $(k \leq N)$ do Steps 3–6:
Gauss-Seidel Iterative Algorithm (2/2)

Step 1  Set $k = 1$
Step 2  While ($k \leq N$) do Steps 3–6:

Step 3  For $i = 1, \ldots, n$

set $x_i = \frac{1}{a_{ii}} \left[ -\sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}XO_j + b_i \right]$
Gauss-Seidel Iterative Algorithm (2/2)

Step 1  Set $k = 1$
Step 2  While ($k \leq N$) do Steps 3–6:

\[
\text{Step 3} \quad \text{For } i = 1, \ldots, n \quad \text{set } x_i = \frac{1}{a_{ii}} \left[ -\sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}XO_j + b_i \right]
\]

Step 4  If $\|x - XO\| < TOL$ then OUTPUT $(x_1, \ldots, x_n)$

(The procedure was successful)

STOP

Step 7  OUTPUT ('Maximum number of iterations exceeded')

STOP

(The procedure was unsuccessful)
Gauss-Seidel Iterative Algorithm (2/2)

Step 1  Set $k = 1$

Step 2  While ($k \leq N$) do Steps 3–6:

Step 3  For $i = 1, \ldots, n$

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

Step 4  If $||x - XO|| < TOL$ then OUTPUT $(x_1, \ldots, x_n)$

(The procedure was successful)

STOP

Step 5  Set $k = k + 1$
Gauss-Seidel Iterative Algorithm (2/2)

Step 1 Set $k = 1$

Step 2 While $(k \leq N)$ do Steps 3–6:

Step 3 For $i = 1, \ldots, n$

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

Step 4 If $||x - X_O|| < TOL$ then OUTPUT $(x_1, \ldots, x_n)$

(The procedure was successful) STOP

Step 5 Set $k = k + 1$

Step 6 For $i = 1, \ldots, n$ set $X_O i = x_i$

 OUTPUT ('Maximum number of iterations exceeded') STOP

(The procedure was unsuccessful)
Gauss-Seidel Iterative Algorithm (2/2)

Step 1  Set $k = 1$

Step 2  While $(k \leq N)$ do Steps 3–6:

Step 3  For $i = 1, \ldots, n$

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

Step 4  If $\|x - XO\| < TOL$ then OUTPUT $(x_1, \ldots, x_n)$

(The procedure was successful)

STOP

Step 5  Set $k = k + 1$

Step 6  For $i = 1, \ldots, n$ set $XO_i = x_i$

Step 7  OUTPUT (‘Maximum number of iterations exceeded’)

STOP  (The procedure was unsuccessful)
Step 3 of the algorithm requires that $a_{ii} \neq 0$, for each $i = 1, 2, \ldots, n$. If one of the $a_{ii}$ entries is 0 and the system is nonsingular, a reordering of the equations can be performed so that no $a_{ii} = 0$. To speed convergence, the equations should be arranged so that $a_{ii}$ is as large as possible. Another possible stopping criterion in Step 4 is to iterate until $\|x(k) - x(k-1)\|/\|x(k)\|$ is smaller than some prescribed tolerance. For this purpose, any convenient norm can be used, the usual being the $l_\infty$ norm.
Step 3 of the algorithm requires that $a_{ii} \neq 0$, for each $i = 1, 2, \ldots, n$. If one of the $a_{ii}$ entries is 0 and the system is nonsingular, a reordering of the equations can be performed so that no $a_{ii} = 0$. To speed convergence, the equations should be arranged so that $a_{ii}$ is as large as possible. Another possible stopping criterion in Step 4 is to iterate until $\|x(k) - x(k-1)\|/\|x(k)\|$ is smaller than some prescribed tolerance. For this purpose, any convenient norm can be used, the usual being the $l_\infty$ norm.
Comments on the Algorithm

- Step 3 of the algorithm requires that $a_{ii} \neq 0$, for each $i = 1, 2, \ldots, n$. If one of the $a_{ii}$ entries is 0 and the system is nonsingular, a reordering of the equations can be performed so that no $a_{ii} = 0$.

- To speed convergence, the equations should be arranged so that $a_{ii}$ is as large as possible.
Gauss-Seidel Iterative Algorithm

Comments on the Algorithm

- Step 3 of the algorithm requires that $a_{ii} \neq 0$, for each $i = 1, 2, \ldots, n$. If one of the $a_{ii}$ entries is 0 and the system is nonsingular, a reordering of the equations can be performed so that no $a_{ii} = 0$.

- To speed convergence, the equations should be arranged so that $a_{ii}$ is as large as possible.

- Another possible stopping criterion in Step 4 is to iterate until

$$\frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|}$$

is smaller than some prescribed tolerance.
Gauss-Seidel Iterative Algorithm

Comments on the Algorithm

- Step 3 of the algorithm requires that $a_{ii} \neq 0$, for each $i = 1, 2, \ldots, n$. If one of the $a_{ii}$ entries is 0 and the system is nonsingular, a reordering of the equations can be performed so that no $a_{ii} = 0$.

- To speed convergence, the equations should be arranged so that $a_{ii}$ is as large as possible.

- Another possible stopping criterion in Step 4 is to iterate until

$$\frac{\|x^{(k)} - x^{(k-1)}\|}{\|x^{(k)}\|}$$

is smaller than some prescribed tolerance.

- For this purpose, any convenient norm can be used, the usual being the $l_\infty$ norm.
Outline

1. The Gauss-Seidel Method
2. The Gauss-Seidel Algorithm
3. Convergence Results for General Iteration Methods
4. Application to the Jacobi & Gauss-Seidel Methods
Introduction

To study the convergence of general iteration techniques, we need to analyze the formula

\[ x^{(k)} = T x^{(k-1)} + c, \quad \text{for each } k = 1, 2, \ldots \]

where \( x^{(0)} \) is arbitrary.

The following lemma and the earlier theorem on convergent matrices provide the key for this study.
Convergence Results for General Iteration Methods

Lemma

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

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

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

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

Proof (1/2)

- Because $Tx = \lambda x$ is true precisely when $(I - T)x = (1 - \lambda)x$, we have $\lambda$ as an eigenvalue of $T$ precisely when $1 - \lambda$ is an eigenvalue of $I - T$. 
Convergence Results for General Iteration Methods

Lemma

If the spectral radius satisfies \( \rho(T) < 1 \), then \((I - T)^{-1}\) exists, and

\[
(I - T)^{-1} = I + T + T^2 + \cdots = \sum_{j=0}^{\infty} T^j
\]

Proof (1/2)

- Because \( T \mathbf{x} = \lambda \mathbf{x} \) is true precisely when \((I - T) \mathbf{x} = (1 - \lambda) \mathbf{x}\), we have \( \lambda \) as an eigenvalue of \( T \) precisely when \( 1 - \lambda \) is an eigenvalue of \( I - T \).

- But \( |\lambda| \leq \rho(T) < 1 \), so \( \lambda = 1 \) is not an eigenvalue of \( T \), and 0 cannot be an eigenvalue of \( I - T \).
Lemma

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

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

Proof (1/2)

- Because $Tx = \lambda x$ is true precisely when $(I - T)x = (1 - \lambda)x$, we have $\lambda$ as an eigenvalue of $T$ precisely when $1 - \lambda$ is an eigenvalue of $I - T$.

- But $|\lambda| \leq \rho(T) < 1$, so $\lambda = 1$ is not an eigenvalue of $T$, and $0$ cannot be an eigenvalue of $I - T$.

- Hence, $(I - T)^{-1}$ exists.
Proof (2/2)

Let

\[ S_m = I + T + T^2 + \cdots + T^m \]
Convergence Results for General Iteration Methods

Proof (2/2)

Let

\[ S_m = I + T + T^2 + \cdots + T^m \]

Then

\[ (I - T)S_m = (1 + T + T^2 + \cdots + T^m) - (T + T^2 + \cdots + T^{m+1}) = I - T^{m+1} \]
Proof (2/2)

Let

\[ S_m = I + T + T^2 + \cdots + T^m \]

Then

\[ (I - T)S_m = (1 + T + T^2 + \cdots + T^m) - (T + T^2 + \cdots + T^{m+1}) = I - T^{m+1} \]

and, since \( T \) is convergent, the Theorem on convergent matrices implies that

\[ \lim_{m \to \infty} (I - T)S_m = \lim_{m \to \infty} (I - T^{m+1}) = I \]
Convergence Results for General Iteration Methods

Proof (2/2)

Let

\[ S_m = I + T + T^2 + \cdots + T^m \]

Then

\[ (I - T)S_m = (1 + T + T^2 + \cdots + T^m) - (T + T^2 + \cdots + T^{m+1}) = I - T^{m+1} \]

and, since \( T \) is convergent, the theorem on convergent matrices implies that

\[ \lim_{m \to \infty} (I - T)S_m = \lim_{m \to \infty} (I - T^{m+1}) = I \]

Thus,

\[ (I - T)^{-1} = \lim_{m \to \infty} S_m = I + T + T^2 + \cdots = \sum_{j=0}^{\infty} T^j \]
Theorem

For any \( x^{(0)} \in \mathbb{R}^n \), the sequence \( \{x^{(k)}\}_{k=0}^{\infty} \) defined by

\[
x^{(k)} = T x^{(k-1)} + c, \quad \text{for each } k \geq 1
\]

converges to the unique solution of

\[
x = T x + c
\]

if and only if \( \rho(T) < 1 \).
Convergence Results for General Iteration Methods

Proof (1/5)

First assume that $\rho(T) < 1$. 

Because $\rho(T) < 1$, the Theorem on convergent matrices implies that $T$ is convergent, and 

$$\lim_{k \to \infty} T^k x(0) = 0$$
First assume that $\rho(T) < 1$. Then,

$$x^{(k)} = Tx^{(k-1)} + c$$
Convergence Results for General Iteration Methods

Proof (1/5)

First assume that \( \rho(T) < 1 \). Then,

\[
\mathbf{x}^{(k)} = T\mathbf{x}^{(k-1)} + \mathbf{c} \\
= T(T\mathbf{x}^{(k-2)} + \mathbf{c}) + \mathbf{c}
\]
Proof (1/5)

First assume that $\rho(T) < 1$. Then,

\[
\begin{align*}
x^{(k)} &= T x^{(k-1)} + c \\
      &= T(T x^{(k-2)} + c) + c \\
      &= T^2 x^{(k-2)} + (T + I)c
\end{align*}
\]
First assume that $\rho(T) < 1$. Then,

\[
\begin{align*}
  x^{(k)} &= T x^{(k-1)} + c \\
  &= T (T x^{(k-2)} + c) + c \\
  &= T^2 x^{(k-2)} + (T + I)c \\
  & \vdots \\
  &= T^k x^{(0)} + (T^{k-1} + \cdots + T + I)c
\end{align*}
\]
Proof (1/5)

First assume that $\rho(T) < 1$. Then,

$$
\begin{align*}
\mathbf{x}^{(k)} &= T \mathbf{x}^{(k-1)} + \mathbf{c} \\
&= T (T \mathbf{x}^{(k-2)} + \mathbf{c}) + \mathbf{c} \\
&= T^2 \mathbf{x}^{(k-2)} + (T + I)\mathbf{c} \\
&\vdots \\
&= T^k \mathbf{x}^{(0)} + (T^{k-1} + \cdots + T + I)\mathbf{c}
\end{align*}
$$

Because $\rho(T) < 1$, the [Theorem] on convergent matrices implies that $T$ is convergent, and

$$
\lim_{k \to \infty} T^k \mathbf{x}^{(0)} = 0
$$
The previous lemma implies that

\[
\lim_{k \to \infty} x^{(k)} = \lim_{k \to \infty} T^k x^{(0)} + \left( \sum_{j=0}^{\infty} T^j \right) c
\]
Convergence Results for General Iteration Methods

Proof (2/5)

The previous lemma implies that

\[
\lim_{k \to \infty} x^{(k)} = \lim_{k \to \infty} T^k x^{(0)} + \left( \sum_{j=0}^{\infty} T^j \right) c
\]

\[
= 0 + (I - T)^{-1} c
\]
The previous lemma implies that

\[
\lim_{k \to \infty} x^{(k)} = \lim_{k \to \infty} T^k x^{(0)} + \left( \sum_{j=0}^{\infty} T^j \right) c
\]

\[
= 0 + (I - T)^{-1} c
\]

\[
= (I - T)^{-1} c
\]
Proof (2/5)

The previous lemma implies that

$$\lim_{k \to \infty} x^{(k)} = \lim_{k \to \infty} T^k x^{(0)} + \left( \sum_{j=0}^{\infty} T^j \right) c$$

$$= 0 + (I - T)^{-1} c$$

$$= (I - T)^{-1} c$$

Hence, the sequence \( \{x^{(k)}\} \) converges to the vector \( x \equiv (I - T)^{-1} c \)
and \( x = Tx + c \).
Proof (3/5)

To prove the converse, we will show that for any \( z \in \mathbb{R}^n \), we have \( \lim_{k \to \infty} T^k z = 0 \).
Convergence Results for General Iteration Methods

Proof (3/5)

- To prove the converse, we will show that for any $z \in \mathbb{R}^n$, we have $\lim_{k \to \infty} T^k z = 0$.
- Again, by the theorem on convergent matrices, this is equivalent to $\rho(T) < 1$. 

Proof (3/5)

To prove the converse, we will show that for any \( z \in \mathbb{R}^n \), we have
\[
\lim_{k \to \infty} T^k z = 0.
\]

Again, by the theorem on convergent matrices, this is equivalent to \( \rho(T) < 1 \).

Let \( z \) be an arbitrary vector, and \( x \) be the unique solution to
\[
x = Tx + c.
\]
Proof (3/5)

To prove the converse, we will show that for any \( z \in \mathbb{R}^n \), we have

\[
\lim_{k \to \infty} T^k z = 0.
\]

Again, by the theorem on convergent matrices, this is equivalent to \( \rho(T) < 1 \).

Let \( z \) be an arbitrary vector, and \( x \) be the unique solution to

\[
x = T x + c.
\]

Define \( x^{(0)} = x - z \), and, for \( k \geq 1 \), \( x^{(k)} = T x^{(k-1)} + c \).
Convergence Results for General Iteration Methods

Proof (3/5)

To prove the converse, we will show that for any \( z \in \mathbb{R}^n \), we have
\[
\lim_{k \to \infty} T^k z = 0.
\]

Again, by the theorem on convergent matrices, this is equivalent to \( \rho(T) < 1 \).

Let \( z \) be an arbitrary vector, and \( x \) be the unique solution to
\[
x = T x + c.
\]

Define \( x^{(0)} = x - z \), and, for \( k \geq 1 \), \( x^{(k)} = T x^{(k-1)} + c \).

Then \( \{x^{(k)}\} \) converges to \( x \).
Convergence Results for General Iteration Methods

Proof (4/5)

Also,

\[ x - x^{(k)} = (Tx + c) - (Tx^{(k-1)} + c) = T(x - x^{(k-1)}) \]
Proof (4/5)

Also,

\[ \mathbf{x} - \mathbf{x}^{(k)} = (T\mathbf{x} + \mathbf{c}) - \left( T\mathbf{x}^{(k-1)} + \mathbf{c} \right) = T \left( \mathbf{x} - \mathbf{x}^{(k-1)} \right) \]

so

\[ \mathbf{x} - \mathbf{x}^{(k)} = T \left( \mathbf{x} - \mathbf{x}^{(k-1)} \right) \]
Proof (4/5)

Also,

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

so

\[ x - x^{(k)} = T \left( x - x^{(k-1)} \right) \]

\[ = T^2 \left( x - x^{(k-2)} \right) \]
Also,

\[ \mathbf{x} - \mathbf{x}^{(k)} = (T \mathbf{x} + \mathbf{c}) - \left( T \mathbf{x}^{(k-1)} + \mathbf{c} \right) = T \left( \mathbf{x} - \mathbf{x}^{(k-1)} \right) \]

so

\[ \mathbf{x} - \mathbf{x}^{(k)} = T \left( \mathbf{x} - \mathbf{x}^{(k-1)} \right) \]

\[ = T^2 \left( \mathbf{x} - \mathbf{x}^{(k-2)} \right) \]

\[ = \vdots \]
Proof (4/5)

Also,

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

so

\[ x - x^{(k)} = T \left( x - x^{(k-1)} \right) = T^2 \left( x - x^{(k-2)} \right) = \cdots = T^k \left( x - x^{(0)} \right) \]
Proof (4/5)

Also,

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

so

\[ x - x^{(k)} = T (x - x^{(k-1)}) \]
\[ = T^2 (x - x^{(k-2)}) \]
\[ = \vdots \]
\[ = T^k (x - x^{(0)}) \]
\[ = T^k z \]
Convergence Results for General Iteration Methods

Proof (5/5)

Hence

$$\lim_{k \to \infty} T^k z = \lim_{k \to \infty} T^k \left( x - x^{(0)} \right)$$
Convergence Results for General Iteration Methods

Proof (5/5)

Hence

\[
\lim_{k \to \infty} T^k z = \lim_{k \to \infty} T^k \left( x - x^{(0)} \right) = \lim_{k \to \infty} \left( x - x^{(k)} \right)
\]
Hence

\[
\lim_{k \to \infty} T^k z = \lim_{k \to \infty} T^k (x - x^{(0)}) \\
= \lim_{k \to \infty} (x - x^{(k)}) \\
= 0
\]
Convergence Results for General Iteration Methods

Proof (5/5)

Hence

\[
\lim_{k \to \infty} T^k z = \lim_{k \to \infty} T^k (x - x^{(0)}) = \lim_{k \to \infty} (x - x^{(k)}) = 0
\]

But \( z \in \mathbb{R}^n \) was arbitrary, so by the theorem on convergent matrices, \( T \) is convergent and \( \rho(T) < 1 \).
Convergence Results for General Iteration Methods

Corollary

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

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

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

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

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

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

(i) \[ \| \mathbf{x} - \mathbf{x}^{(k)} \| \leq \| T \|^k \| \mathbf{x}^{(0)} - \mathbf{x} \| \]
Corollary

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

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

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

(i) \( \| \mathbf{x} - \mathbf{x}^{(k)} \| \leq \| T \|^k \| \mathbf{x}^{(0)} - \mathbf{x} \| \)

(ii) \( \| \mathbf{x} - \mathbf{x}^{(k)} \| \leq \frac{\| T \|^k}{1 - \| T \|} \| \mathbf{x}^{(1)} - \mathbf{x}^{(0)} \| \)
Convergence Results for General Iteration Methods

**Corollary**

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

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

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

(i) \[ \| \mathbf{x} - \mathbf{x}^{(k)} \| \leq \| T \|^k \| \mathbf{x}^{(0)} - \mathbf{x} \| \]

(ii) \[ \| \mathbf{x} - \mathbf{x}^{(k)} \| \leq \frac{\| T \|^k}{1 - \| T \|} \| \mathbf{x}^{(1)} - \mathbf{x}^{(0)} \| \]

The proof of the following corollary is similar to that for the Corollary to the Fixed-Point Theorem for a single nonlinear equation.
Outline

1. The Gauss-Seidel Method
2. The Gauss-Seidel Algorithm
3. Convergence Results for General Iteration Methods
4. Application to the Jacobi & Gauss-Seidel Methods
We have seen that the Jacobi and Gauss-Seidel iterative techniques can be written

\[ x^{(k)} = T_j x^{(k-1)} + c_j \]  and 
\[ x^{(k)} = T_g x^{(k-1)} + c_g \]

where \( T_j \) and \( T_g \) are defined as:

\[ T_j = D - 1(L + U) \]  and 
\[ T_g = (D - L)^{-1} U \]

If \( \rho(T_j) \) or \( \rho(T_g) \) is less than 1, then the corresponding sequence \( \{x^{(k)}\}_{k=0}^{\infty} \) will converge to the solution \( x \) of \( Ax = b \).
Using the Matrix Formulations

We have seen that the Jacobi and Gauss-Seidel iterative techniques can be written

\[
\begin{align*}
x^{(k)} &= T_j x^{(k-1)} + c_j \quad \text{and} \\
x^{(k)} &= T_g x^{(k-1)} + c_g
\end{align*}
\]

using the matrices

\[
T_j = D^{-1}(L + U) \quad \text{and} \quad T_g = (D - L)^{-1}U
\]

respectively.
Convergence of the Jacobi & Gauss-Seidel Methods

Using the Matrix Formulations

We have seen that the Jacobi and Gauss-Seidel iterative techniques can be written

\[ x^{(k)} = T_j x^{(k-1)} + c_j \quad \text{and} \quad x^{(k)} = T_g x^{(k-1)} + c_g \]

using the matrices

\[ T_j = D^{-1}(L + U) \quad \text{and} \quad T_g = (D - L)^{-1}U \]

respectively. If \( \rho(T_j) \) or \( \rho(T_g) \) is less than 1, then the corresponding sequence \( \{x^{(k)}\}_{k=0}^{\infty} \) will converge to the solution \( x \) of \( Ax = b \).
Convergence of the Jacobi & Gauss-Seidel Methods

Example

For example, the Jacobi method has

\[ x^{(k)} = D^{-1}(L + U)x^{(k-1)} + D^{-1}b, \]
Convergence of the Jacobi & Gauss-Seidel Methods

Example

For example, the Jacobi method has

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

and, if \( \{x^{(k)}\}_{k=0}^{\infty} \) converges to \( x \),

Since \( D^{-1}L - U = A \), the solution \( x \) satisfies \( Ax = b \).
Example

For example, the Jacobi method has

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

and, if $\{x^{(k)}\}_{k=0}^{\infty}$ converges to $x$, then

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

For example, the Jacobi method has

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

and, if \( \{x^{(k)}\}_{k=0}^{\infty} \) converges to \( x \), then

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

This implies that

\[ Dx = (L + U)x + b \quad \text{and} \quad (D - L - U)x = b \]
For example, the Jacobi method has

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

and, if \( \{\mathbf{x}^{(k)}\}_{k=0}^\infty \) converges to \( \mathbf{x} \), then

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

This implies that

\[ D\mathbf{x} = (L + U)\mathbf{x} + \mathbf{b} \quad \text{and} \quad (D - L - U)\mathbf{x} = \mathbf{b} \]

Since \( D - L - U = A \), the solution \( \mathbf{x} \) satisfies \( A\mathbf{x} = \mathbf{b} \).
The following are easily verified sufficiency conditions for convergence of the Jacobi and Gauss-Seidel methods.
Convergence of the Jacobi & Gauss-Seidel Methods

The following are easily verified sufficiency conditions for convergence of the Jacobi and Gauss-Seidel methods.

**Theorem**

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 converge to the unique solution of $Ax = b$. 
Convergence of the Jacobi & Gauss-Seidel Methods

Is Gauss-Seidel better than Jacobi?

No general results exist to tell which of the two techniques, Jacobi or Gauss-Seidel, will be most successful for an arbitrary linear system. In special cases, however, the answer is known, as is demonstrated in the following theorem.
Is Gauss-Seidel better than Jacobi?

- No general results exist to tell which of the two techniques, Jacobi or Gauss-Seidel, will be most successful for an arbitrary linear system.
Is Gauss-Seidel better than Jacobi?

- No general results exist to tell which of the two techniques, Jacobi or Gauss-Seidel, will be most successful for an arbitrary linear system.

- In special cases, however, the answer is known, as is demonstrated in the following theorem.
Convergence of the Jacobi & Gauss-Seidel Methods

(Stein-Rosenberg) Theorem

If $a_{ij} \leq 0$, for each $i \neq j$ and $a_{ii} > 0$, for each $i = 1, 2, \ldots, n$, then one and only one of the 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) $\rho(T_j) = \rho(T_g) = 1$

(Stein-Rosenberg) Theorem

If \( a_{ij} \leq 0 \), for each \( i \neq j \) and \( a_{ii} > 0 \), for each \( i = 1, 2, \ldots, n \), then one and only one of the 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) \( \rho(T_j) = \rho(T_g) = 1 \)

For the proof of this result, see pp. 120–127. of

Two Comments on the Theorem

For the special case described in the theorem, we see from part (i), namely

\[ 0 \leq \rho(T_g) < \rho(T_j) < 1 \]
Two Comments on the Theorem

For the special case described in the theorem, we see from part (i), namely

$$0 \leq \rho(T_g) < \rho(T_j) < 1$$

that when one method gives convergence, then both give convergence,
Two Comments on the Theorem

For the special case described in the theorem, we see from part (i), namely

\[ 0 \leq \rho(T_g) < \rho(T_j) < 1 \]

that when one method gives convergence, then both give convergence, and the Gauss-Seidel method converges faster than the Jacobi method.
Two Comments on the Theorem

- For the special case described in the theorem, we see from part (i), namely
  \[ 0 \leq \rho(T_g) < \rho(T_j) < 1 \]
  that when one method gives convergence, then both give convergence, and the Gauss-Seidel method converges faster than the Jacobi method.

- Part (ii), namely
  \[ 1 < \rho(T_j) < \rho(T_g) \]
  indicates that when one method diverges then both diverge, and the divergence is more pronounced for the Gauss-Seidel method.
Questions?
Theorem

The following statements are equivalent.

(i) \( A \) is a convergent matrix.

(ii) \( \lim_{n \to \infty} \|A^n\| = 0 \), for some natural norm.

(iii) \( \lim_{n \to \infty} \|A^n\| = 0 \), for all natural norms.

(iv) \( \rho(A) < 1 \).

(v) \( \lim_{n \to \infty} A^n x = 0 \), for every \( x \).

Fixed-Point Theorem

Let $g \in C[a, b]$ be such that $g(x) \in [a, b]$, for all $x$ in $[a, b]$. Suppose, in addition, that $g'$ exists on $(a, b)$ and that a constant $0 < k < 1$ exists with

$$|g'(x)| \leq k, \quad \text{for all } x \in (a, b).$$

Then for any number $p_0$ in $[a, b]$, the sequence defined by

$$p_n = g(p_{n-1}), \quad n \geq 1$$

converges to the unique fixed point $p$ in $[a, b]$. 

Return to the Corollary to the Fixed-Point Theorem
Corollary to the Fixed-Point Convergence Result

If $g$ satisfies the hypothesis of the Fixed-Point Theorem then

$$|p_n - p| \leq \frac{k^n}{1 - k} |p_1 - p_0|$$

Return to the Corollary to the Convergence Theorem for General Iterative Methods