## Example

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

$$\begin{aligned}
10x_1 - x_2 + 2x_3 \quad &= 6,\\
-x_1 + 11x_2 - x_3 + 3x_4 &= 25,\\
2x_1 - x_2 + 10x_3 - x_4 &= -11,\\
3x_2 - x_3 + 8x_4 &= 15
\end{aligned}$$

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

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

## *Solution*

Previously, this problem was solved by the Jacobi method with 10 iterations.

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

$$
\begin{aligned}
x_1^{(k)} &= \qquad\qquad \frac{1}{10}x_2^{(k-1)} - \frac{1}{5}x_3^{(k-1)} \qquad\qquad + \frac{3}{5},\\
x_2^{(k)} &= \frac{1}{11}x_1^{(k)} \qquad\qquad + \frac{1}{11}x_3^{(k-1)} - \frac{3}{11}x_4^{(k-1)} + \frac{25}{11}\\
x_3^{(k)} &= -\frac{1}{5}x_1^{(k)} + \frac{1}{10}x_2^{(k)} \qquad\qquad + \frac{1}{10}x_4^{(k-1)} - \frac{11}{10}\\
x_4^{(k)} &= \qquad\qquad - \frac{3}{8}x_2^{(k)} + \frac{1}{8}x_3^{(k)} \qquad\qquad + \frac{15}{8}
\end{aligned}
$$

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 value in the below table:

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $x_1^{(k)}$ | 0.0000 | 0.6000 | 1.030 | 1.0065 | 1.0009 | 1.0001 |
| $x_2^{(k)}$ | 0.0000 | 2.3272 | 2.037 | 2.0036 | 2.0003 | 2.0000 |
| $x_3^{(k)}$ | 0.0000 | −0.9873 | −1.014 | −1.0025 | −1.0003 | −1.0000 |
| $x_4^{(k)}$ | 0.0000 | 0.8789 | 0.9844 | 0.9983 | 0.9999 | 1.0000 |

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.

Note: Jacobi method requires twice as many iterations for the same accuracy.

## Writing the Gauss-Seidel method in matrix form:

Multiplying both sides of equation,

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

By $a_{ii}$ and collecting all *k*th iterate terms, results in:

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

for each $i = 1, 2, \ldots, n$. Writing all $n$ equations gives

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

with the definitions of $D$, $L$, and $U$ given previously, we have

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

and

$$\mathbf{x}^{(k)} = (D - L)^{-1}U\mathbf{x}^{(k-1)} + (D - L)^{-1}\mathbf{b}, \quad \text{for each } k = 1, 2, \ldots$$

Letting $T_g = (D-L)^{-1}U$ and $\mathbf{c}_g = (D-L)^{-1}\mathbf{b}$, gives

$$\mathbf{x}^{(k)} = T_g \mathbf{x}^{(k-1)} + \mathbf{c}_g$$

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

## Eigenvalues and Eigenvectors

If $A$ is a square matrix, the **characteristic polynomial** of $A$ is defined by

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

- $\lambda$ is an eigenvalue of $A$ if and only if $\det(A - \lambda I) = 0$.
- corresponding eigenvector $\mathbf{x} \neq \mathbf{0}$ is determined by solving the system

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

## Example

Determine the eigenvalues and eigenvectors for the matrix

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

***Solution*** The characteristic polynomial of $A$ is

$$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 0 & 0 \\ 1 & 1-\lambda & 2 \\ 1 & -1 & 4-\lambda \end{bmatrix}$$

$$= -(\lambda^3 - 7\lambda^2 + 16\lambda - 12) = -(\lambda - 3)(\lambda - 2)^2,$$

so there are two eigenvalues of $A$: $\lambda_1 = 3$ and $\lambda_2 = 2$.

An eigenvector $\mathbf{x}_1$ corresponding to the eigenvalue $\lambda_1 = 3$ is a solution to the equation $(A - 3 \cdot I)\mathbf{x}_1 = \mathbf{0}$, so

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -2 & 2 \\ 1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

which implies that $x_1 = 0$ and $x_2 = x_3$. So,

$$\mathbf{x}_1 = (0, 1, 1)^t$$

Similarly,

$$\lambda_2 = 2 \quad \Rightarrow \quad (A - 2 \cdot I)\mathbf{x} = \mathbf{0} \quad \Rightarrow$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 2 \\ 1 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

In this case the eigenvector has only to satisfy the equation

$$x_1 - x_2 + 2x_3 = 0$$

which can be done in various ways. For example,

$$x_1 = 0 \Rightarrow x_2 = 2x_3 \Rightarrow \mathbf{x}_2 = (0, 2, 1)^t$$

$$x_2 = 0 \Rightarrow x_1 = -2x_3 \Rightarrow \mathbf{x}_3 = (-2, 0, 1)^t$$

## Spectral Radius

The **spectral radius** $\rho(A)$ of a matrix $A$ is defined by

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

## *Theorem*

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

## Theorem

If $\rho(T) < 1$ 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

$$\|\mathbf{x}^{(k)} - \mathbf{x}\| \approx \rho(T)^k \|\mathbf{x}^{(0)} - \mathbf{x}\|$$

# Relaxation Techniques for solving Linear Systems

Relaxation Techniques are used to accelerate Convergence.

## Definition

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

Suppose that $x^{(k)}$ is an approximate solution defined by:

$$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_{i-1}^{(k)}, x_i^{(k-1)}, \ldots, x_n^{(k-1)})^t$$

The corresponding residual vector is,

$$r^{(k)} = b - Ax^{(k)}$$

In particular, the *i*th component of residual vector is,

$$r_i^{(k)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k-1)} - a_{ii} x_i^{(k-1)}$$

So,

$$a_{ii} x_i^{(k-1)} + r_i^{(k)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k-1)} \qquad (7.14)$$

Recall, however, that in the Gauss-Seidel method, $x_i^{(k)}$ is chosen to be

$$x_i^{(k)} = \frac{1}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)}\right] \tag{7.15}$$

so Eq. (7.14) can be rewritten as

$$a_{ii}x_i^{(k-1)} + r_i^{(k)} = a_{ii}x_i^{(k)}$$

Consequently,

$$x_i^{(k)} = x_i^{(k-1)} + \frac{r_i^{(k)}}{a_{ii}} \tag{7.16}$$

If we modify the above equation as,

$$x_i^{(k)} = x_i^{(k-1)} + \omega\frac{r_i^{(k)}}{a_{ii}} \tag{7.17}$$

then for certain choices of positive $\omega$ we can obtain significantly faster convergence.

Methods involving Eq. (7.17) are called **relaxation methods**.

$0 < \omega < 1$ ➡ **under-relaxation methods**

$1 < \omega$ ➡ **over-relaxation methods**

For the Gauss-Seidel method, we are interested in choices of $\omega > 1$.

This method is Called SOR (Successive Over-Relaxation) method.

by using Eq. (7.14), we can reformulate Eq. (7.17) for calculation purposes as

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

Combination of the above equation and equation (7.15) results in,

$$x_i^{(k)} = x_i^{(k-1)} + \omega(x_i^{(k)} - x_i^{(k-1)})$$

Thus, at the end of each iteration of the Gauss-Seidel method $x_i^{(k)}$ is modified using the above equation.

To determine the matrix form of the SOR method, we rewrite equation (*) as,

$$a_{ii}x_i^{(k)} + \omega \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} = (1-\omega)a_{ii}x_i^{(k-1)} - \omega \sum_{j=i+1}^{n} a_{ij}x_j^{(k-1)} + \omega b_i$$

so that in vector form, we have

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

That is,

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

Letting $T_\omega = (D - \omega L)^{-1}[(1-\omega)D + \omega U]$ and $\mathbf{c}_\omega = \omega(D - \omega L)^{-1}\mathbf{b}$,

$$\mathbf{x}^{(k)} = T_\omega \mathbf{x}^{(k-1)} + \mathbf{c}_\omega$$

## Example

The linear system $A\mathbf{x} = \mathbf{b}$ given by

$$\begin{aligned} 4x_1 + 3x_2 \qquad\quad &= 24, \\ 3x_1 + 4x_2 - \phantom{4}x_3 &= 30, \\ -\phantom{4}x_2 + 4x_3 &= -24, \end{aligned}$$

has the solution $(3, 4, -5)^t$. Compare the iterations from the Gauss-Seidel method and the SOR method with $\omega = 1.25$ using $\mathbf{x}^{(0)} = (1, 1, 1)^t$ for both methods.

***Solution*** For each $k = 1, 2, \ldots,$ the equations for the Gauss-Seidel method are

$$\begin{aligned} x_1^{(k)} &= -0.75x_2^{(k-1)} + 6, \\ x_2^{(k)} &= -0.75x_1^{(k)} + 0.25x_3^{(k-1)} + 7.5, \\ x_3^{(k)} &= 0.25x_2^{(k)} - 6, \end{aligned}$$

For the SOR method, the Gauss-Seidel equations are modified as,

$$x_1^{(k)} = x_1^{(k-1)} + 1.25\left(x_1^{(k)} - x_1^{(k-1)}\right)$$

$$x_2^{(k)} = x_2^{(k-1)} + 1.25\left(x_2^{(k)} - x_2^{(k-1)}\right)$$

$$x_3^{(k)} = x_3^{(k-1)} + 1.25\left(x_3^{(k)} - x_3^{(k-1)}\right)$$

The first seven iterates for each method are listed in Tables 7.3 and 7.4.

**Table 7.3**

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $x_1^{(k)}$ | 1 | 5.250000 | 3.1406250 | 3.0878906 | 3.0549316 | 3.0343323 | 3.0214577 | 3.0134110 |
| $x_2^{(k)}$ | 1 | 3.812500 | 3.8828125 | 3.9267578 | 3.9542236 | 3.9713898 | 3.9821186 | 3.9888241 |
| $x_3^{(k)}$ | 1 | −5.046875 | −5.0292969 | −5.0183105 | −5.0114441 | −5.0071526 | −5.0044703 | −5.0027940 |

**Table 7.4**

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $x_1^{(k)}$ | 1 | 6.312500 | 2.6223145 | 3.1333027 | 2.9570512 | 3.0037211 | 2.9963276 | 3.0000498 |
| $x_2^{(k)}$ | 1 | 3.5195313 | 3.9585266 | 4.0102646 | 4.0074838 | 4.0029250 | 4.0009262 | 4.0002586 |
| $x_3^{(k)}$ | 1 | −6.6501465 | −4.6004238 | −5.0966863 | −4.9734897 | −5.0057135 | −4.9982822 | −5.0003486 |

For the iterates to be accurate to seven decimal places, the Gauss-Seidel method requires 34 iterations, as opposed to 14 iterations for the SOR method with $\omega = 1.25$.