1 Iterative Methods: Gauss-Seidel Method (8.1-8.3)

The Gauss-Seidel Method is another iterative method that is based on the Jacobi method discussed earlier. For any matrix A, we can write it into three parts:

$$A = D + L + U.$$

Observing a 4×4 matrix, this might look like

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix}}_{D} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & a_{32} & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 0 \end{bmatrix}}_{L} + \underbrace{\begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

(Example.) Suppose

$$A = \begin{bmatrix} 3 & 0 & -8 \\ -2 & -1 & 7 \\ 1 & 5 & 2 \end{bmatrix}.$$

If we decompose this into $D + L + U$, we have

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 1 & 5 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} 0 & 0 & -8 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{bmatrix}.$$

1.1 Deriving from Jacobi Method

Recall that we defined the iterative process for the Jacobi method

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j \right) \qquad i = 1, 2, \ldots, n \text{ and } a_{ii} \neq 0.$$

With this in mind, notice that

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

Note that $x_j^{(k+1)}$, where $1 \leq j \leq i-1$, is already known, so we can just use that. Thus, we have

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

the Gauss-Seidel Method. Representing and restructuring this in matrix form, we have

$$\begin{aligned} x^{(k+1)} &= D^{-1}(b - Lx^{(k+1)} - Ux^{(k)}) \\ &\implies Dx^{(k+1)} = b - Lx^{(k+1)} - Ux^{(k)} \\ &\implies Dx^{(k+1)} + Lx^{(k+1)} = b - Ux^{(k)} \\ &\implies (D+L)x^{(k+1)} = b - Ux^{(k)} \\ &\implies \boxed{x^{(k+1)} = (D+L)^{-1}(b - Ux^{(k)})} \end{aligned}$$

Remarks:

- $D + L$ is the lower part of A. It is invertible if and only if all diagonal entries (a_{ii}) are non-zero.
- $D + L = A - U$. Thus, you can re-represent the Gauss-Seidel method as

$$x^{(k+1)} = (A-U)^{-1}(b - Ux^{(k)}).$$

(Example.) Let $A = \begin{bmatrix} 3 & 0 & -8 \\ -2 & -1 & 7 \\ 1 & 5 & 2 \end{bmatrix}$, $x^* = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, and $Ax^* = b = \begin{bmatrix} -5 \\ 4 \\ 8 \end{bmatrix}$. Suppose we have the initial guess $x^{(0)} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$. Using the formula for each entry, we have the following example:

- Starting at $k = 0$, for our first iteration, we have

$$\begin{aligned} x_1^{(1)} &= \frac{1}{a_{11}} \left(b_1 - \sum_{j=1}^{1-1} a_{1j} x_j^{(0+1)} - \sum_{j=1+1}^{3} a_{1j} x_j^{(0)} \right) \\ &= \frac{1}{a_{11}} \left(b_1 - \sum_{j=1}^{0} a_{1j} x_j^{(1)} - \sum_{j=2}^{3} a_{1j} x_j^{(0)} \right) \\ &= \frac{1}{a_{11}} \left(b_1 - \sum_{j=2}^{3} a_{1j} x_j^{(0)} \right) \\ &= \frac{1}{a_{11}} \left(b_1 - (a_{12} x_2^{(0)} + a_{13} x_3^{(0)}) \right). \end{aligned}$$

$$\begin{aligned} x_2^{(1)} &= \frac{1}{a_{22}} \left(b_2 - \sum_{j=1}^{2-1} a_{2j} x_j^{(0+1)} - \sum_{j=2+1}^{3} a_{2j} x_j^{(0)} \right) \\ &= \frac{1}{a_{22}} \left(b_2 - \sum_{j=1}^{1} a_{2j} x_j^{(1)} - \sum_{j=3}^{3} a_{2j} x_j^{(0)} \right) \\ &= \frac{1}{a_{22}} \left(b_2 - a_{21} x_1^{(1)} - (a_{23} x_3^{(0)}) \right). \end{aligned}$$

$$
\begin{aligned}
x_3^{(1)} &= \frac{1}{a_{33}}\left(b_3 - \sum_{j=1}^{3-1} a_{3j}x_j^{(0+1)} - \sum_{j=3+1}^{3} a_{3j}x_j^{(0)}\right) \\
&= \frac{1}{a_{33}}\left(b_3 - \sum_{j=1}^{2} a_{3j}x_j^{(1)} - \sum_{j=4}^{3} a_{3j}x_j^{(0)}\right) \\
&= \frac{1}{a_{33}}\left(b_3 - \sum_{j=1}^{2} a_{3j}x_j^{(1)}\right) \\
&= \frac{1}{a_{33}}\left(b_3 - (a_{31}x_1^{(1)} + a_{32}x_2^{(1)})\right).
\end{aligned}
$$

This gives us $x^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix}$. The remaining iterations have the same idea.

Alternatively, using the matrix-based formula:

- Starting at $k = 0$, for our first iteration, we have

$$x^{(1)} = (D + L)^{-1}(b - Ux^{(0)}).$$

This requires us to find $(D + L)^{-1}$. Once we find this, we can perform the remaining iterations.

1.2 Convergence Result

The Guass-Seidel method converges for every initial guess $x^{(0)}$ if all eigenvalues λ of $G = I - (D+L)^{-1}A = I - (A - U)^{-1}A$ satisfy the following:

- All eigenvalues are distinct.
- $|\lambda| < 1$.

Proof. The proof is similar to the proof for Jacobi. We want to show that $e^{(k)} = x^* - x^{(k)} \mapsto 0$.

$$
\begin{aligned}
e^{(k+1)} &= x^* - x^{(k+1)} \\
&= x^* - (D + L)^{-1}(b - (A - L - D)x^{(k)}) \\
&= x^* - (D + L)^{-1}(b - Ax^{(k)} - (L + D)x^{(k)}) \\
&= x^* - (D + L)^{-1}(b - Ax^{(k)}) - (D + L)^{-1}((L + D)x^{(k)}) \\
&= x^* - (D + L)^{-1}(b - Ax^{(k)}) - x^{(k)} \\
&= x^* - (D + L)^{-1}(b - A(x^* - e^{(k)})) - x^{(k)} \\
&= x^* - (D + L)^{-1}(b - Ax^* + Ae^{(k)}) - x^{(k)} \\
&= x^* - (D + L)^{-1}(b - Ax^* + A(x^* - x^{(k)})) - x^{(k)} \\
&= x^* - (D + L)^{-1}(\underbrace{b - Ax^*}_{0} + Ax^* - Ax^{(k)}) - x^{(k)} \\
&= x^* - (D + L)^{-1}(Ax^* - Ax^{(k)}) - x^{(k)} \\
&= x^* - (D + L)^{-1}(A(x^* - x^{(k)})) - x^{(k)} \\
&= x^* - (D + L)^{-1}(Ae^{(k)}) - x^{(k)} \\
&= x^* - x^{(k)} - (D + L)^{-1}(Ae^{(k)}) \\
&= e^{(k)} - (D + L)^{-1}(Ae^{(k)}) \\
&= (I - (D + L)^{-1}A)e^{(k)}
\end{aligned}
$$

Overall, $e^{(k+1)} = (I-(D+L)^{-1}A)e^{(k)} = (I-(D+L)^{-1}A)^2e^{(k-1)} = \ldots = (I-(D+L)^{-1}A)^{k+1}e^{(0)}$. So,

$$e^{(k)} = (I-(D+L)^{-1}A)^k e^{(0)}.$$

As $G = I-(D+L)^{-1}A$, we end up with

$$e^{(k)} = G^k e^{(0)}.$$

Since all eigenvalues are distinct, this means that G has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$. Let $\lambda_1, \ldots, \lambda_n$ denote the corresponding eigenvalues of G. Then,

$$Gv_i = \lambda_i v_i \qquad i = 1, 2, \ldots, n.$$

As $\{v_i\}$ for $i = 1, 2, \ldots, n$ are linearly independent, we can express $e^{(0)}$ as a linear combination of them,

$$e^{(0)} = c_1v_1 + c_2v_2 + \ldots + c_nv_n.$$

Then, the rest of the operation works similarly to the power method. That is,

$$Ge^{(0)} = c_1\lambda_1v_1 + c_2\lambda_2v_2 + \ldots + c_n\lambda_nv_n.$$

$$G^2e^{(0)} = c_1\lambda_1^2v_1 + c_2\lambda_2^2v_2 + \ldots + c_n\lambda_n^2v_n.$$

$$G^ke^{(0)} = c_1\lambda_1^kv_1 + c_2\lambda_2^kv_2 + \ldots + c_n\lambda_n^kv_n.$$

Consequently,

$$||e^{(k)}|| = ||G^ke^{(0)}|| \leq |c_1||\lambda_1^k|||v_1|| + |c_2||\lambda_2^k|||v_2|| + \ldots + |c_n||\lambda_n^k|||v_n||.$$

Since $|\lambda_i^k| \mapsto 0$ if and only if $|\lambda_i| < 1$, we have $||e^{(k)}|| \mapsto 0$ for every initial guess $x^{(0)}$ if all eigenvalues $|\lambda| < 1$. $\square$