Gauss-Jacobi and Gauss Seidel Methods

The Gauss-Jacobi and Gauss Seidel methods are related and similar iterative algorithms used to solve systems of linear equations, particularly useful for large systems where expensive.

Given a system of linear equations:

$$A\,\underline{x} = \underline{b}\,,$$

where A is a square matrix of coefficients, $\underline{x}$ is the vector of unknowns, and $\underline{b}$ is the vector of constants, the Jacobi method solves for $\underline{x}$ iteratively.

The method involves splitting the matrix A into the sum of a diagonal matrix D and the remainder, R :

$$A = D + R\,.$$

Hence

$$(D + R\,)\,\underline{x} = \underline{b}\,,$$

and therefore

$$D\,\underline{x} = \underline{b} - R\,\underline{x}\,,$$

The methods involve assigning initial values to the components of $\underline{x}$, $\underline{x}^{(0)}$ and then using this formula to iteratively update to $\underline{x}^{(1)}, \underline{x}^{(2)}, \underline{x}^{(3)}$ ... until the estimate for $\underline{x}$ is sufficiently close to the exact values for $\underline{x}$. The exact value for $\underline{x}$ may be unknown, in which case, if an iteration makes a small enough change in the estimate for $\underline{x}$ then the method could halt.

The methods are only guaranteed to converge to $\underline{x}$ if the matrix A is *diagonally dominant*; on each row the absolute value of the diagonal component is greater than the sum of the absolute values of the off-diagonal components:

$$|a_{ii}| > \sum_{i \neq j} |a_{ij}|\,.$$

The initial value or ‘guess’ $\underline{x}^{(0)}$ can be anything and, provided we have diagonal dominance in A, convergence to $\underline{x}$ is guaranteed. However, less iterations, and therefore less effort is required if the $\underline{x}^{(0)}$ is close to $\underline{x}$. Given the diagonal dominance of A, probably the best initial estimate for $\underline{x}$ is through removing the remainder term from the equation above:

$$D\,\underline{x}^{(0)} = \underline{b}\,;$$

each component of $\underline{x}^{(0)}$ is simply the corresponding component of $\underline{b}$ divided by the diagonal component of A; $x_i^{(0)} = b_i / a_{ii}$.

In order to demonstrate the methods, let us consider the 3×3 system:

$$\begin{pmatrix} -4 & 2 & 1 \\ -1 & 10 & -4 \\ 1 & -2 & 5 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -9 \\ -16 \\ 9 \end{pmatrix},$$

in which the matrix is diagonally dominant. The system is in the form $A\,\underline{x} = \underline{b}$ is which,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} -4 & 2 & 1 \\ -1 & 10 & -4 \\ 1 & -2 & 5 \end{pmatrix}, \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -9 \\ -16 \\ 9 \end{pmatrix} \text{and } \underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$$

To demonstrate the methods, the solution $\underline{x}$ is presumed unknown and the problem is to find successive approximations to $\underline{x}$ using an iterative method.

$$\begin{pmatrix} -4 & 2 & 1 \\ -1 & 10 & -4 \\ 1 & -2 & 5 \end{pmatrix} \underline{x} = \begin{pmatrix} -9 \\ -16 \\ 9 \end{pmatrix}.$$

As stated above, the matrix is split into two, one matrix consisting of the diagonal components and the other consisting of the remaining off-diagonal components:

$$\left[\begin{pmatrix} -4 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 5 \end{pmatrix} + \begin{pmatrix} 0 & 2 & 1 \\ -1 & 0 & -4 \\ 1 & -2 & 0 \end{pmatrix}\right] \underline{x} = \begin{pmatrix} -9 \\ -16 \\ 9 \end{pmatrix}.$$

Putting the remainder matrix on the other side gives the following:

$$\begin{pmatrix} -4 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 5 \end{pmatrix} \underline{x} = \begin{pmatrix} -9 \\ -16 \\ 9 \end{pmatrix} - \begin{pmatrix} 0 & 2 & 1 \\ -1 & 0 & -4 \\ 1 & -2 & 0 \end{pmatrix} \underline{x}\,.$$

Although not necessarily part of these methods, as stated earlier, a good initial estimate for $\underline{x}$ is obtained by solving the above system without the remainder

$$\begin{pmatrix} -4 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 5 \end{pmatrix} \underline{x}^{(0)} = \begin{pmatrix} -9 \\ -16 \\ 9 \end{pmatrix},$$

so that $\underline{x}^{(0)} = \begin{pmatrix} 2.25 \\ -1.6 \\ 1.8 \end{pmatrix}$.

The Gauss-Jacobi and Gauss-Seidel methods are stated in the following sections and applied to the example problem for two iterations. A spreadsheet[1] continues to twelve iterations for this example. The spreadsheet also allows the user to change the 3×3 system and gives twelve iterations of the two methods.

[1] Gauss Jacobi and Gauss Seidel methods for a 3x3 system (Spreadsheet)

Gauss-Jacobi Iteration

The Gauss-Jacobi method updates the solution iteratively as follows:

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

Hence $x_1^{(1)} = \frac{1}{-4}(-9 - 2 \times (-1.6) - 1 \times 1.8) = 1.9$, $x_2^{(1)} = \frac{1}{10}(-16 - (-1) \times 2.25 - (-4) \times 1.8) = -0.655$ and $x_3^{(1)} = \frac{1}{5}(9 - 1 \times 2.25 - (-2) \times (-1.6)) = 0.71$. So that $\underline{x}^{(1)} = \begin{pmatrix} 1.9 \\ -0.655 \\ 0.71 \end{pmatrix}$.

For the second iteration, $x_1^{(2)} = \frac{1}{-4}(-9 - 2 \times (-0.655) - 1 \times 0.71) = 2.1$, $x_2^{(2)} = \frac{1}{10}(-16 - (-1) \times 1.9 - (-4) \times 0.71) = -1.126$ and $x_3^{(2)} = \frac{1}{5}(9 - 1 \times 2.25 - (-2) \times (-1.6)) = 1.158$. So that $\underline{x}^{(2)} = \begin{pmatrix} 2.1 \\ -1.126 \\ 1.158 \end{pmatrix}$.

Gauss-Seidel Iteration

The Gauss-Jacobi method updates the solution iteratively as follows:

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

Put simply, as $x_1^{(k+1)}, x_2^{(k+1)}$... are evaluated at the $k+1$st step, they are used as the most up-to-date and therefore probably their best-available estimate, rather than the previous estimates $x_1^{(k)}, x_2^{(k)}$... that are used in the Gauss-Jacobi method

Hence $x_1^{(1)} = \frac{1}{-4}(-9 - 2 \times (-1.6) - 1 \times 1.8) = 1.9$, $x_2^{(1)} = \frac{1}{10}(-16 - (-1) \times 1.9 - (-4) \times 1.8) = -0.69$ and $x_3^{(1)} = \frac{1}{5}(9 - 1 \times 1.9 - (-2) \times (-0.69)) = 1.144$. So that $\underline{x}^{(1)} = \begin{pmatrix} 1.9 \\ -0.69 \\ 1.144 \end{pmatrix}$.

For the second iteration, $x_1^{(2)} = \frac{1}{-4}(-9 - 2 \times (-0.69) - 1 \times 1.144) = 2.191$, $x_2^{(2)} = \frac{1}{10}(-16 - (-1) \times 2.191 - (-4) \times 1.144) = -0.9233$ and $x_3^{(2)} = \frac{1}{5}(9 - 1 \times 2.191 - (-2) \times (-0.9233)) = 0.92248$. So that $\underline{x}^{(2)} = \begin{pmatrix} 2.191 \\ -0.9233 \\ 0.92248 \end{pmatrix}$.