# Iterative Methods

The following summarizes the main points of our class discussion of the classical iterative methods for solving $Ax = b$ and also provides additional useful results.

Conventions: The system dimensions are $A \in \mathbb{R}^{n\times n}$ and $x \in \mathbb{R}^n$, $b \in \mathbb{R}^n$. The $ij$th entry in $A$ is denoted by $a_{ij}$. The $i$th components of $x$ and $b$ are denoted by $x_i$ and $b_i$. A sum is omitted when its lower index of summation is greater than its upper index of summation.

**The classical methods.**

These methods are based on a "splitting" $A = D - L - U$, in which $D$ is a diagonal matrix containing the diagonal of $A$, and $-L$ and $-U$ are strict lower- and upper-triangular matrices containing the strict lower- and upper-triangular parts of $A$. The "matrix" forms of the methods are as follows:

<u>JACOBI ITERATION:</u>

Given an initial $x$,

Iterate:

$$x \leftarrow D^{-1}\left[(L+U)x + b\right]$$

<u>GAUSS–SEIDEL ITERATION:</u>

Given an initial $x$,

Iterate:

$$x \leftarrow (D-L)^{-1}\big(Ux + b\big)$$

<u>SUCCESSIVE OVER-RELAXATION (SOR):</u>

Given an initial $x$,

Iterate:

$$x \leftarrow (D-\omega L)^{-1}\left\{\left[(1-\omega)D + \omega U\right]x + \omega b\right\}$$

The equivalent "componentwise" forms of the methods are as follows

<u>JACOBI ITERATION:</u>

Given an initial $x$,

Iterate:

For $i = 1, \ldots, n$

$$x_i^+ = \Big(b_i - \textstyle\sum_{j\neq i} a_{ij}x_j\Big)/a_{ii}$$

Update $x \leftarrow x^+$.

<u>GAUSS–SEIDEL ITERATION:</u>

Given an initial $x$,

Iterate:

For $i = 1, \ldots, n$

$$x_i \leftarrow \Big(b_i - \textstyle\sum_{j<i} a_{ij}x_j - \sum_{j>i} a_{ij}x_j\Big)/a_{ii}$$

Successive Over-Relaxation (SOR):

Given an initial $x$,

Iterate:

For $i = 1, \ldots, n$

$$x_i \leftarrow (1-\omega)x_i + (\omega/a_{ii})\left(b_i - \sum_{j<i} a_{ij}x_j - \sum_{j>i} a_{ij}x_j\right)$$

**Convergence theory.**

Consider the following General Iteration, in which $T \in \mathbb{R}^{n\times n}$ and $c \in \mathbb{R}^n$:

General Iteration:

Given an initial $x$,

Iterate:

$x \leftarrow Tx + c$

The classical iterative methods are of this form, as follows:

- Jacobi iteration: $T = T_{\mathrm{J}} \equiv D^{-1}(L+U)$ and $c = c_{\mathrm{J}} \equiv D^{-1}b$.
- Gauss–Seidel iteration: $T = T_{\mathrm{GS}} \equiv (D-L)^{-1}U$ and $c = c_{\mathrm{GS}} \equiv (D-L)^{-1}b$.
- SOR: $T = T_\omega \equiv (D-\omega L)^{-1}\left[(1-\omega)D + \omega U\right]$ and $c = c_\omega \equiv \omega(D-\omega L)^{-1}b$.

Note that for all three methods, $x^* = Tx^* + c$ if and only if $x^* = A^{-1}b$. Thus, if the iterates produced by one of these methods converge, then they converge to $A^{-1}b$.

Proposition 1 and Theorem 3 below are results for the General Iteration.

Proposition 1: *If $\{x^{(k)}\}$ produced by the General Iteration converges to some $x^*$, then $x^* = Tx^* + c$.*

Definition 2: *The* spectrum *and the* spectral radius *of $T$ are, respectively,*

$$\sigma(T) \equiv \{\lambda : Tx = \lambda x, \text{ for some } x \neq 0\} \quad \textit{and} \quad \rho(T) \equiv \max_{\lambda\in\sigma(T)} |\lambda|.$$

Theorem 3: *The iterates $\{x^{(k)}\}$ produced by the General Iteration converge for every $x^{(0)}$ if and only if $\rho(T) < 1$. If $\rho(T) < 1$, then for every $x^{(0)}$, $\{x^{(k)}\}$ converges to the unique $x^*$ satisfying $x^* = Tx^* + c$.*

The results below pertain to convergence of the classical iterations and are often useful in applications.

Definition 4: *$A$ is* diagonally dominant *if $|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|$ for $1 \leq i \leq n$. $A$ is* strictly diagonally dominant *if $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$ for $1 \leq i \leq n$.*

Theorem 5: *If $A$ is strictly diagonally dominant, then $A$ is nonsingular. Moreover, $\rho(T_{\mathrm{J}}) < 1$ and $\rho(T_{\mathrm{GS}}) < 1$; consequently, both the Jacobi and Gauss–Seidel iterates converge to $A^{-1}b$ for every $x^{(0)}$.*

THEOREM 6 (STEIN–ROSENBERG): *If $a_{ij} \leq 0$ for $i \neq j$ and if $a_{ii} > 0$ for each $i$, then one and only one of the following holds:*

(a) $0 \leq \rho(T_{GS}) < \rho(T_J) < 1$, (b) $1 < \rho(T_J) < \rho(T_{GS})$,
(c) $\rho(T_J) = \rho(T_{GS}) = 0$, (d) $\rho(T_J) = \rho(T_{GS}) = 1$.

Note that if (a) holds, then both the Jacobi and Gauss–Seidel iterates converge to $A^{-1}b$ for every $x^{(0)}$, and we can expect the Gauss–Seidel iterates to converge faster. If (c) holds, then the iterates from both methods reach $A^{-1}b$ in a finite number of iterations. If (b) or (d) holds, then the iterates do not converge for some $x^{(0)}$.

THEOREM 7: *If $A$ is symmetric positive-definite* (SPD)*, then the Gauss–Seidel iterates converge to $A^{-1}b$ for every $x^{(0)}$.*

Note that, since an SPD matrix has positive diagonal elements, it follows from Theorem 7 and the Stein–Rosenberg theorem that if $A$ is SPD with non-positive off-diagonal elements, then the Jacobi iterates as well as the Gauss–Seidel iterates converge to $A^{-1}b$ for every $x^{(0)}$, and we can expect the Gauss–Seidel iterates to converge faster.

THEOREM 8 (KAHAN): *If $a_{ii} \neq 0$ for $i = 1, \ldots, n$, then the* SOR *iteration matrix $T_\omega$ satisfies $\rho(T_\omega) \geq |\omega - 1|$. Consequently, the* SOR *iterates converge for every $x^{(0)}$ only if $0 < \omega < 2$.*

THEOREM 9 (OSTROWSKI–REICH): *If $A$ is symmetric positive-definite and $0 < \omega < 2$, then the* SOR *iterates converge to $A^{-1}b$ for every $x^{(0)}$.*

THEOREM 10: *If $A$ is symmetric positive-definite and tridiagonal, then $\rho(T_{\mathrm{GS}}) = \rho(T_{\mathrm{J}})^2 < 1$, and the $\omega$ that minimizes $\rho(T_\omega)$ is*

$$\omega = \frac{2}{1 + \sqrt{1 - \rho(T_{\mathrm{J}})^2}}.$$

*For this $\omega$, $\rho(T_\omega) = \omega - 1$.*

The results above came from reference [1, Sec. 7.3], although Theorem 5 has been augmented a bit and Theorem 7 is not stated there, presumably because it is implied by Theorem 9 with $\omega = 1$. There are many more convergence results for the classical iterations. A good general reference is [2]. Seminal classical references are [3] and [4].

**References.**


1. R. L. Burden and J. D. Faires, *Numerical Analysis* (9th ed.), Thomson–Brooks/Cole, 2010..
2. J. Stoer and R. Bullirsch, *Introduction to Numerical Analysis*, Springer-Verlag, 1980.
3. R. S. Varga, *Matrix Iterative Analysis*, Series in Automatic Computation, Prentice–Hall, 1962.
4. D. M. Young, *Iterative Solution of Large Linear Systems*, Computer Science and Applied Mathematics, Academic Press, 1971.