Lecture 6: Iterative methods

Michael S. Floater

September 16, 2025

These notes are based on Sections 2.11–2.12 of Chapter 2 of the book.

1 Iterative schemes and splitting

Given a linear system $A\mathbf{x} = \mathbf{b}$ where A is an $n \times n$ matrix and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$, solving it by factorization can be expensive for large n. It might be more efficient to use an iterative method. The simplest iterative schemes are based on so-called *splitting*. We choose a matrix B and rewrite the system as

$$(A - B)\mathbf{x} = -B\mathbf{x} + \mathbf{b}. \tag{1}$$

We choose B such that $A - B$ is non-singular and such that the system

$$(A - B)\mathbf{x} = \mathbf{y}$$

is easy to solve for any right-hand side $\mathbf{y}$. For example, we might choose B to make $A - B$ diagonal or triangular.

Having chosen B we use an iterative method to find $\mathbf{x}$. We make an initial guess (approximation) $\mathbf{x}^{(0)}$ for the solution $\mathbf{x}$ although $\mathbf{x}^{(0)}$ can be arbitrary. We then generate a sequence $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$, $\mathbf{x}^{(3)}$, and so on, by solving

$$(A - B)\mathbf{x}^{(k+1)} = -B\mathbf{x}^{(k)} + \mathbf{b}, \qquad k = 0, 1, 2, \ldots. \tag{2}$$

If this sequence converges to a limit

$$\mathbf{x} = \lim_{k\to\infty} \mathbf{x}^{(k)},$$

then by taking the limit of both sides of (2), we obtain equation (1) and therefore $\mathbf{x}$ solves $A\mathbf{x} = \mathbf{b}$.

What are the necessary and sufficient conditions for convergence? Suppose that A is non-singular so that there is a unique solution $\mathbf{x}$. Consider the k-th error,

$$\mathbf{e}^{(k)} := \mathbf{x}^{(k)} - \mathbf{x}.$$

By subtracting equation (1) from equation (2), we deduce that

$$(A - B)\mathbf{e}^{(k+1)} = -B\mathbf{e}^{(k)}, \qquad k = 0, 1, 2, \ldots.$$

Under our assumption that $A - B$ is non-singular this means that

$$\mathbf{e}^{(k+1)} = H\mathbf{e}^{(k)}, \qquad k = 0, 1, 2, \ldots, \tag{3}$$

where the matrix

$$H := -(A - B)^{-1}B$$

is the *iteration matrix*. In practical applications we do not calculate H. We are just using it here to analyze the convergence of (2). Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues (real or complex) of H. Recall that the *spectral radius* of H is

$$\rho(H) = \max_{i=1,\ldots,n} |\lambda_i|.$$

Theorem 1 *The error vectors $\mathbf{e}^{(k)}$ converge to zero as $k \to \infty$ if and only if $\rho(H) < 1$.*

Proof. Applying (3) recursively gives

$$\mathbf{e}^{(k)} = H^k\mathbf{e}^{(0)}.$$

Thus the theorem is equivalent to the fact that $H^k \to 0$ as $k \to \infty$ if and only if $\rho(H) < 1$. This can be proved using the Jordan normal form of H (see R. Varga, *Matrix iterative analysis*). □

2 Jacobi and Gauss-Seidel iterations

Both of these splitting methods can be used when A has non-zero diagonal elements. We write A in the form $A = L + D + U$ where L is the strictly lower triangular (subdiagonal) part of A, D is the diagonal, and U is the strictly upper triangular (superdiagonal) part of A.

2.1 Jacobi iteration

We choose $B = L + U$, so that $A - B = D$, the diagonal part of A. Then

$$D\mathbf{x}^{(k+1)} = -(L+U)\mathbf{x}^{(k)} + \mathbf{b}, \qquad k = 0, 1, 2, \ldots.$$

Written out in full,

$$x_i^{(k+1)} = \frac{1}{A_{i,i}} \left(-\sum_{j \neq i} A_{i,j} x_j^{(k)} + b_i \right), \qquad i = 1, \ldots, n.$$

2.2 Gauss-Seidel iteration

We choose $B = U$, so that $A - B = L + D$, which is lower triangular. Then

$$(L+D)\mathbf{x}^{(k+1)} = -U\mathbf{x}^{(k)} + \mathbf{b}, \qquad k = 0, 1, 2, \ldots.$$

There is no need to invert $L + D$: we just use forward substitution,

$$x_i^{(k+1)} = \frac{1}{A_{i,i}} \left(-\sum_{j<i} A_{i,j} x_j^{(k+1)} - \sum_{j>i} A_{i,j} x_j^{(k)} + b_i \right), \qquad i = 1, \ldots, n.$$

An advantage of Gauss-Seidel iteration compared to Jacobi iteration is that in the implementation we do not need to keep both of the vectors $\mathbf{x}^{(k)}$ and $\mathbf{x}^{(k+1)}$. We just maintain a single vector $\mathbf{y}$ and update it, replacing y_i by

$$\frac{1}{A_{i,i}} \left(-\sum_{j \neq i} A_{i,j} y_j + b_i \right)$$

in sequence for $i = 1, \ldots, n$.

Sufficient conditions for these two iterations to converge are as follows. A matrix A is said to be *strictly diagonally dominant* if

$$|A_{i,i}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |A_{i,j}| \qquad \text{for } i = 1, \ldots, n.$$

Strictly diagonally dominant matrices are non-singular (due to the Gerschgorin theorem).

Theorem 2 *If A is strictly diagonally dominant, then both the Jacobi and the Gauss-Seidel methods converge.*

Theorem 3 *If A is symmetric and positive definite, then the Gauss-Seidel method converges. If A is symmetric and both A and $2D - A$ are positive definite, then the Jacobi method converges.*

We will just go through the proof of the first one.

Proof of Theorem 2. For the Jacobi method we need to show that the eigenvalues of

$$H = -D^{-1}(L + U)$$

are less than one in absolute value. Suppose λ is an eigenvalue of H. Then

$$\det(H - \lambda I) = 0.$$

Since A is strictly diagonally dominant, D is non-singular and we can multiply by $\det(D)$ so that

$$\det(L + U + \lambda D) = 0.$$

Suppose that $|\lambda| \geq 1$. Then since A is strictly diagonally dominant, so is $L + U + \lambda D$ which is therefore non-singular and its determinant cannot be zero. We thus have a contradiction and we conclude that $|\lambda| < 1$ as required.

For the Gauss-Seidel method we need to show that the eigenvalues of

$$H = -(L + D)^{-1}U$$

are less than one in absolute value. We suppose that

$$\det(H - \lambda I) = 0.$$

Since $L + D$ is non-singular, we can multiply by $\det(L + D)$ so that

$$\det(U + \lambda(L + D)) = 0.$$

Suppose that $|\lambda| \geq 1$. Then since A is strictly diagonally dominant, so is $U + \lambda(L + D)$ which is therefore non-singular and its determinant cannot be zero. This is a contradiction and we conclude that $|\lambda| < 1$. $\square$

3 Exercises

Exercise 2.10 *Consider the iteration* $\mathbf{e}^{(k+1)} = H\mathbf{e}^{(k)}$ *for* $k = 0, 1, 2, \ldots$, *where*

$$H = \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix},$$

with $\alpha, \beta, \gamma \in \mathbb{R}$ *and* γ *large and* $|\alpha| < 1$, $|\beta| < 1$. *Calculate* H^k *and show that its elements tend to zero as* $k \to \infty$. *Hence deduce that* $\mathbf{e}^{(k)} \to 0$ *as* $k \to \infty$.

Exercise 2.11 *Starting with an arbitrary* $\mathbf{x}^{(0)}$ *the sequence* $\mathbf{x}^{(k)}$, $k = 1, 2, \ldots$, *is calculated by*

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}^{(k+1)} + \begin{bmatrix} 0 & 0 & 0 \\ \alpha & 0 & 0 \\ \gamma & \beta & 0 \end{bmatrix} \mathbf{x}^{(k)} = \mathbf{b}$$

in order to solve the linear system

$$\begin{bmatrix} 1 & 1 & 1 \\ \alpha & 1 & 1 \\ \gamma & \beta & 1 \end{bmatrix} \mathbf{x} = \mathbf{b},$$

where α, β, γ *are constants. Find all values for* α, β, γ *such that the sequence converges for every* $\mathbf{x}^{(0)}$ *and* $\mathbf{b}$. *What happens when* $\alpha = \beta = \gamma = -1$ *and when* $\alpha = \beta = 0$*?*