Iterative methods

Consider the linear system

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

Iterative methods start from an initial guess $\underline{x}^{(0)}$ and construct a sequence of approximate solutions $\{\underline{x}^{(k)}\}$ such that

$$\underline{x} = \lim_{k \to \infty} \underline{x}^{(k)}.$$

Splitting methods

The matrix A is split as

$$A = M - N$$

Splitting methods go like

$$\underline{x}^{(0)} \text{ given} \quad \text{solve} \quad M\underline{x}^{(k)} = \underline{b} + N\underline{x}^{(k-1)} \quad k = 1, 2, \cdots \tag{1}$$

With iterative methods we give up the idea of computing the exact solution, but we want low operational costs. In particular:

- the system (1) must be much easier to deal with than the original system $Ax = b$, that is, the matrix M must be as simple as possible, and of course non-singular;
- the sequence $\{\underline{x}^{(k)}\}$ must converge to $\underline{x}$ for any initial guess $\underline{x}^{(0)}$;
- the convergence must be fast.

Different choices for M give rise to different iterative methods.

Jacobi method

take $M = diag(A)$ (and hence $N = M - A$), applicable if $a_{ii} \neq 0 \ \forall i$. At each iteration k we have to solve a diagonal system

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ 0 & \cdots & \ddots & 0 \\ 0 & \cdots & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} - \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ \vdots \\ x_n^{(k-1)} \end{pmatrix}$$

Thus we obtain

$$\boxed{x_i^{(k)} = \Big(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k-1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k-1)}\Big) / a_{ii} \quad i = 1, \cdots, n}$$

The number of operations for each component is $\sim 2n$, so that the cost for one Jacobi iteration is $\sim 2n^2$.

Gauss-Seidel method

take $M = tril(A)$, applicable if $a_{ii} \neq 0 \; \forall i$. At each iteration k we have to solve a lower triangular system

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} - \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ \vdots \\ x_n^{(k-1)} \end{pmatrix}$$

Thus we obtain

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

The difference with respect to Jacobi method is in the first sum of the formula, where the updated $x_j^{(k)}$ are used instead of the old $x_j^{(k-1)}$. The number of operations is exactly the same: for each component is $\sim 2n$, so that the cost for one Gauss-Seidel iteration is $\sim 2n^2$.

Convergence analysis for splitting methods

In all cases we want convergence for any initial guess $\underline{x}^{(0)}$. With paper and pencil we study the error at each iteration.
Let $e^{(k)} = \underline{x} - \underline{x}^{(k)}$ be the error at the k^{th} iteration.
Since $\underline{x}$ and $\underline{x}^{(k)}$ are solutions of

$$M\underline{x} = \underline{b} + N\underline{x}, \qquad M\underline{x}^{(k)} = \underline{b} + N\underline{x}^{(k-1)},$$

by subtracting we get

$$M(\underline{x} - \underline{x}^{(k)}) = N(\underline{x} - \underline{x}^{(k-1)}) \implies e^{(k)} = \underbrace{M^{-1}N}_{B} e^{(k-1)}$$

where $B = M^{-1}N$ is the iteration matrix.

$$e^{(k)} = Be^{(k-1)} \quad k = 1, 2, \cdots, \implies e^{(k)} = B^k e^{(0)}.$$

If we want $\lim_{k\to\infty} e^{(k)} = 0$ we need $\lim_{k\to\infty} B^k = 0$.

Convergent matrices

A matrix $B \in \mathbb{R}^{n \times n}$ is convergent if

$$\lim_{k \to \infty} B^k = 0,$$

where 0 is the matrix identically zero. Then:

Lemma 1

Let $B \in \mathbb{R}^{n \times n}$. We have

$$\lim_{k \to \infty} B^k = 0 \iff \max_i |\lambda_i(B)| < 1.$$

The proof is not trivial for a generic B.

A useful property of natural norm of matrices

Lemma 2

Let $|||A|||$ be any natural norm of matrix. Then

$$\max_i |\lambda_i(A)| \leq |||A||| \quad \forall A \in \mathbb{R}^{n\times n}.$$

Proof.

Let λ be an eigenvalue of A, and let $\underline{v} \neq 0$ an eigenvector associated to λ, that is $A\underline{v} = \lambda\underline{v}$. From the properties of the norms we immediately have

$$|\lambda|\|\underline{v}\| = \|\lambda\underline{v}\| = \|A\underline{v}\| \leq |||A|||\|\underline{v}\|,$$

then $|\lambda|\|\underline{v}\| \leq |||A|||\|\underline{v}\|$, and then $|\lambda| \leq |||A|||$. □

The quantity $\max_i |\lambda_i(A)|$ is called the spectral radius of A, and denoted as $\rho(A)$.

The matrix $|||\cdot|||_\infty$ norm

Lemma 3

Given $B \in \mathbb{R}^{n\times n}$, *the natural norm* $|||B|||_\infty := \sup_{v\in\mathbb{R}^n} \frac{\|Bv\|_\infty}{\|v\|_\infty}$ *can be rewritten as*

$$|||B|||_\infty = \max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}|$$

Proof.

Let us start by proving that $|||B|||_\infty \leq \max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}|$. It holds:

$$\|Bv\|_\infty = \max_{i=1,\dots,n} |(Bv)_i| = \max_{i=1,\dots,n} |\sum_{j=1}^{n} B_{i,j} v_j| \leq \max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j} v_j|$$

$$\leq \max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}||v_j| \leq \|v\|_\infty \max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}|$$

continue . . .

Proof.

Therefore, for all $v \in \mathbb{R}^n$, it holds

$$\frac{\|Bv\|_\infty}{\|v\|_\infty} \leq \max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}|$$

and finally

$$|||B|||_\infty = \sup_{v \in \mathbb{R}^n} \frac{\|Bv\|_\infty}{\|v\|_\infty} \leq \max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}|$$

It remains to prove that $\max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}| \leq |||B|||_\infty$. Let $\widehat{i}$ be the row index that realizes the maximum and let us define $w \in \mathbb{R}^n$ as $w_j = \mathrm{sign}(B_{\widehat{i},j})$. We observe that $\|w\|_\infty = 1$.

continue ...

Proof.

It holds

$$\begin{aligned}
\max_{i=1,\dots,n} \sum_{j=1}^{n} |B_{i,j}| &= \sum_{j=1}^{n} |B_{\hat{i},j}| = \sum_{j=1}^{n} B_{\hat{i},j} w_j = |\sum_{j=1}^{n} B_{\hat{i},j} w_j| \\
&\leq \max_{i=1,\dots,n} |\sum_{j=1}^{n} B_{i,j} w_j| = \|Bw\|_\infty = \frac{\|Bw\|_\infty}{\|w\|_\infty} \\
&\leq \sup_{v \in \mathbb{R}^n} \frac{\|Bv\|_\infty}{\|v\|_\infty} = |||B|||_\infty
\end{aligned}$$

□

Classes of matrices for which we have convergence results

Lemma 4

If A is diagonally dominant, i.e.,

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \qquad \forall i = 1, 2, \cdots, n$$

both Jacobi and Gauss-Seidel converge.

Proof.

We shall prove the Lemma only for Jacobi method. The iteration matrix B_J is given by

$$B_J = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \cdots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \cdots & 0 \end{bmatrix}$$

Since A is diagonally dominant, $|||B_J|||_\infty = \max_i \sum_{j \neq i} |\frac{a_{ij}}{a_{ii}}| < 1$, and we deduce (from Lemma 1) that $\max_i |\lambda_i(B_J)| < 1$. Then B_J is convergent and Jacobi method converges. □

Lemma 5

If A is symmetric and positive definite Gauss-Seidel converges. Jacobi might or might not converge.