Systems of linear equations

Iterative methods

Systems of linear equations

- Direct methods
- Iterative methods

Large-scale systems of linear equations

Take $\dim \mathbf{A} = m \sim 10^6$

Memory

$$m^2 = 10^{12} \text{ elements} \sim 10^{11} \text{ bytes} \sim 10^2 \text{ Gb}$$

Flops

$$\sim m^3 = 10^{18} \text{ operations}$$

At 1 GFlop/sec

$$\sim 10^9 \text{ sec} \sim 10 \cdots 10^2 \text{ yrs}$$

Sparse linear algebra

In practice, matrices are often *sparse*: only a small fraction of elements is non-zero.

- Partial differential equations: grid discretizations
- Adjacency matrices of large graphs and networks
- ...

Want to only reference non-zero elements

Sparse linear algebra

Want to only reference non-zero elements and avoid fill-ins.

Fill-in: a matrix element which initally is zero, but becomes non-zero during the execution of an algorithm.

Direct methods (generally) generate many fill-ins.

Iterative methods for systems of linear equations

Given a system of equations

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

with $\mathbf{A}$ an $m \times m$ matrix, $\det \mathbf{A} \neq 0$

Identically rewrite it:

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{c}$$

Iterative methods for systems of linear equations

Given a system of equations

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

with $\mathbf{A}$ an $m \times m$ matrix, $\det \mathbf{A} \neq 0$

Identically rewrite it:

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{c}$$

One simple way: $\mathbf{D} = \operatorname{diag} \mathbf{A}$ and

$$\mathbf{A} = \mathbf{D} + (\mathbf{A} - \mathbf{D})$$

Iterative methods for systems of linear equations

The original system of equations,

$$\mathbf{D}\mathbf{x} + (\mathbf{A} - \mathbf{D})\mathbf{x} = \mathbf{b}$$

becomes

$$\mathbf{x} = \mathbf{D}^{-1}(\mathbf{D} - \mathbf{A})\mathbf{x} + \mathbf{D}^{-1}\mathbf{b}$$

$$B_{jl} = \begin{cases} 0\,, & j = l \\ -a_{jl}/a_{jj}\,, & j \neq l \end{cases}$$

Clearly, need $a_{jj} \neq 0, \quad j = 1, \ldots, m$

Jacobi iteration

Start with

$$\mathbf{x}^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_m^{(0)} \end{pmatrix}$$

And iterate ($k = 0, 1, \ldots$)

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{c}$$

Jacobi iteration

Start with

$$\mathbf{x}^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_m^{(0)} \end{pmatrix}$$

And iterate ($k = 0, 1, \ldots$)

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{c}$$

- Does it converge?
- What is the rate of convergence?
- *A priori* and *a posteriori* error bounds?

Jacobi iteration, convergence

Define $\widehat{\mathbf{x}}$ as the solution of $\widehat{\mathbf{x}} = \mathbf{B}\widehat{\mathbf{x}} + \mathbf{c}$.
Let $\|\mathbf{B}\| < 1$.

Then,

1. $\mathbf{x}^{(n)}$ converges to $\widehat{\mathbf{x}}$: $\quad \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \to 0$ as $n \to \infty$
2. $\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \leqslant \|\mathbf{B}\|^n \|\mathbf{x}^{(0)} - \widehat{\mathbf{x}}\|$

Jacobi iteration, convergence

$$\left.\begin{aligned} \widehat{\mathbf{x}} &= \mathbf{B}\widehat{\mathbf{x}} + \mathbf{c} \\ \mathbf{x}^{(n)} &= \mathbf{B}\mathbf{x}^{(n-1)} + \mathbf{c} \end{aligned}\right| \quad \Rightarrow \mathbf{x}^{(n)} - \widehat{\mathbf{x}} = \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right)$$

Therefore,

$$\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \leqslant \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\|$$

Jacobi iteration, convergence

$$\left.\begin{aligned} \widehat{\mathbf{x}} &= \mathbf{B}\widehat{\mathbf{x}} + \mathbf{c} \\ \mathbf{x}^{(n)} &= \mathbf{B}\mathbf{x}^{(n-1)} + \mathbf{c} \end{aligned}\right| \quad \Rightarrow \mathbf{x}^{(n)} - \widehat{\mathbf{x}} = \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right)$$

Therefore,

$$\begin{aligned} \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| &\leqslant \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\| \\ &\leqslant \|\mathbf{B}\|^2 \cdot \|\mathbf{x}^{(n-2)} - \widehat{\mathbf{x}}\| \end{aligned}$$

Jacobi iteration, convergence

$$\left.\begin{aligned} \widehat{\mathbf{x}} &= \mathbf{B}\widehat{\mathbf{x}} + \mathbf{c} \\ \mathbf{x}^{(n)} &= \mathbf{B}\mathbf{x}^{(n-1)} + \mathbf{c} \end{aligned}\right| \quad \Rightarrow \mathbf{x}^{(n)} - \widehat{\mathbf{x}} = \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right)$$

Therefore,

$$\begin{aligned} \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| &\leqslant \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\| \\ &\leqslant \|\mathbf{B}\|^2 \cdot \|\mathbf{x}^{(n-2)} - \widehat{\mathbf{x}}\| \leqslant \cdots \leqslant \|\mathbf{B}\|^n \cdot \|\mathbf{x}^{(0)} - \widehat{\mathbf{x}}\| \end{aligned}$$

So that we have established statement 2.

Jacobi iteration, convergence

$$\left.\begin{aligned}\widehat{\mathbf{x}} &= \mathbf{B}\widehat{\mathbf{x}} + \mathbf{c} \\ \mathbf{x}^{(n)} &= \mathbf{B}\mathbf{x}^{(n-1)} + \mathbf{c}\end{aligned}\right| \quad \Rightarrow \mathbf{x}^{(n)} - \widehat{\mathbf{x}} = \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right)$$

Therefore,

$$\begin{aligned}\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| &\leqslant \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\| \\ &\leqslant \|\mathbf{B}\|^2 \cdot \|\mathbf{x}^{(n-2)} - \widehat{\mathbf{x}}\| \leqslant \cdots \leqslant \|\mathbf{B}\|^n \cdot \|\mathbf{x}^{(0)} - \widehat{\mathbf{x}}\|\end{aligned}$$

So that we have established statement 2.

Convergence follows: $\|\mathbf{B}\| \leqslant 1 \quad \Rightarrow \qquad \|\mathbf{B}\|^n \to 0$ as $n \to \infty$
and

$$\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \to 0 \quad \text{as} \quad n \to \infty$$

Jacobi iteration, *a posteriori* error bound

We have an *a priori error* bound. What about an *a posteriori* one?

Jacobi iteration, *a posteriori* error bound

We have an *a priori error* bound. What about an *a posteriori* one?

If $\|\mathbf{B}\| < 1$, then

$$\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \leqslant q \cdot \|\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}\|,$$

with

$$q = \frac{\|\mathbf{B}\|}{1 - \|\mathbf{B}\|}$$

Jacobi iteration, *a posteriori* error bound

$$
\begin{aligned}
\mathbf{x}^{(n)} - \widehat{\mathbf{x}} &= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right) \\
&= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)}\right) + \mathbf{B}\left(\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\right)
\end{aligned}
$$

Jacobi iteration, *a posteriori* error bound

$$\begin{aligned}
\mathbf{x}^{(n)} - \widehat{\mathbf{x}} &= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right) \\
&= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)}\right) + \mathbf{B}\left(\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\right)
\end{aligned}$$

Using the triangle inequality,

$$\Rightarrow \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \leqslant \|\mathbf{B}(\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)})\| + \|\mathbf{B}(\mathbf{x}^{(n)} - \widehat{\mathbf{x}})\|$$

Jacobi iteration, *a posteriori* error bound

$$
\begin{aligned}
\mathbf{x}^{(n)} - \widehat{\mathbf{x}} &= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right) \\
&= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)}\right) + \mathbf{B}\left(\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\right)
\end{aligned}
$$

Using the triangle inequality,

$$
\begin{aligned}
\Rightarrow \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| &\leqslant \|\mathbf{B}(\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)})\| + \|\mathbf{B}(\mathbf{x}^{(n)} - \widehat{\mathbf{x}})\| \\
&\leqslant \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)}\| + \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\|
\end{aligned}
$$

Jacobi iteration, *a posteriori* error bound

$$\begin{aligned}\mathbf{x}^{(n)} - \widehat{\mathbf{x}} &= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \widehat{\mathbf{x}}\right) \\ &= \mathbf{B}\left(\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)}\right) + \mathbf{B}\left(\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\right)\end{aligned}$$

Using the triangle inequality,

$$\begin{aligned}\Rightarrow \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| &\leqslant \|\mathbf{B}(\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)})\| + \|\mathbf{B}(\mathbf{x}^{(n)} - \widehat{\mathbf{x}})\| \\ &\leqslant \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n-1)} - \mathbf{x}^{(n)}\| + \|\mathbf{B}\| \cdot \|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\|\end{aligned}$$

And indeed,

$$\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \leqslant \frac{\|\mathbf{B}\|}{1 - \|\mathbf{B}\|} \|\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}\|$$

Recap: Simple iteration for systems of linear equations

Given

$$\mathbf{A}\mathbf{x} = \mathbf{b} \qquad \iff \qquad \mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{c}$$

Jacobi iteration

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{c}$$

Sufficient condition for convergence: $\|\mathbf{B}\| < 1$

$$\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \leqslant \|\mathbf{B}\|^n \|\mathbf{x}^{(0)} - \widehat{\mathbf{x}}\|$$

and

$$\|\mathbf{x}^{(n)} - \widehat{\mathbf{x}}\| \leqslant \frac{\|\mathbf{B}\|}{1 - \|\mathbf{B}\|} \|\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}\|$$

Convergence rate of simple iteration

Consider the ∞-norm:

$$\|\mathbf{B}\|_\infty = \max_{1 \leqslant j \leqslant m} \underbrace{\sum_{l=1}^{m} |B_{jl}|}_{\text{sum over the } j\text{-th row}} < 1$$

Recall that

$$B_{jj} = 0 \qquad \text{and} \qquad B_{jl} = -a_{jl}/a_{jj}$$

Convergence rate of simple iteration

The condition $\|\mathbf{B}\|_\infty < 1$ is equivalent to $\mathbf{A}$ being *diagonally dominant*:

$$\sum_{l \neq j} |a_{jl}| < |a_{jj}|$$

for $j = 1, \ldots, m$

Loosely speaking, Jacobi iteration converges well if $\mathbf{A}$ is "close" to a diagonal matrix.

Necessary and sufficient conditions for convergence

Let $\lambda_1, \dots, \lambda_m$ are eigenvalues of $\mathbf{B}$.

The *spectral radius* of $\mathbf{B}$,

$$\rho(\mathbf{B}) = \max_{1 \leqslant j \leqslant m} |\lambda_j|$$

Simple iterations converge if and only if

$$\rho(\mathbf{B}) < 1$$

Necessary and sufficient conditions for convergence

Let $\lambda_1, \dots, \lambda_m$ are eigenvalues of $\mathbf{B}$.

The *spectral radius* of $\mathbf{B}$,

$$\rho(\mathbf{B}) = \max_{1 \leqslant j \leqslant m} |\lambda_j|$$

Simple iterations converge if and only if

$$\rho(\mathbf{B}) < 1$$

If $\mathbf{B}$ is symmetric, $\rho(\mathbf{B}) \equiv \|\mathbf{B}\|_2$