1   Iterative Methods: Jacobi Method (8.1)

We now return to the problem from the beginning of the class: solving $A\mathbf{x} = \mathbf{b}$, for $A \in \mathbb{R}^{n \times n}$ and $\mathbf{b} \in \mathbb{R}^n$. Recall that we performed Gauss elimination directly, resulting in a flop count of $\mathcal{O}(n^3)$. The problem with this is that, if n is large (hundreds of thousands of equations), then Gauss elimination may not be feasible.

Aside from Gaussian Elimination, there are iterative methods that we can run to solve $A\mathbf{x} = \mathbf{b}$. The idea is that we need an initial guess $\mathbf{x}^{(0)}$. Then, from $\mathbf{x}^{(k)}$, we hope to somehow iterate to $\mathbf{x}^{(k+1)}$. The idea is that, as $k \mapsto \infty$, $\mathbf{x}^{(k)} \mapsto \mathbf{x}^*$, where $\mathbf{x}^*$ is the solution to $A\mathbf{x} = \mathbf{b}$ (i.e., $A\mathbf{x}^* = \mathbf{b}$). Some advantages of this process include

- If $\mathbf{x}^{(0)}$ is already close to $\mathbf{x}^*$, then the iterative method converges *fast*.

- We can stop at any point of the iteration, depending on how accurate the approximation of $\mathbf{x}^*$ (by $\mathbf{x}^{(k)}$) should be.

1.1   The Jacobi Method

The idea is that ith equation in the system $A\mathbf{x} = \mathbf{b}$ is

$$\sum_{j=1}^{n} a_{ij} x_j = b_i.$$

This can be rewritten by solving for x_i and assuming that $a_{ii} \neq 0$ as

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j \right) \qquad i = 1, 2, \ldots, n.$$

The iteration process, then, is

$$\boxed{x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)} \right)} \qquad i = 1, 2, \ldots, n.$$

As mentioned, this is an iterative process, so we **stop** when $||\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}|| < \epsilon$, where we can choose any vector norm.

1.2   Matrix Representation

We can also represent this process using matrices. In equation form, we have

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)} \right) \qquad i = 1, 2, \ldots, n.$$

Notice that, for $i = 1, 2, \ldots, n$, $\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)}$ is the same thing as saying

$$\sum_{i=1}^{n} a_{ij} x_j^{(k)} - a_{ii} x_i^{(k)} \implies A\mathbf{x}^{(k)} - D\mathbf{x}^{(k)},$$

where

$$D = \begin{bmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & \ddots & & \\ & & & & a_{nn} \end{bmatrix}.$$

Therefore, translating our equation into matrix form means translating each component into matrix form. This gives us

$$\mathbf{x}^{(k+1)} = \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{a_{11}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{a_{33}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{a_{nn}} \end{bmatrix}}_{D^{-1}} \left(\mathbf{b} - (A - D)\mathbf{x}^{(k)} \right).$$

This further gives us, for $k \geq 0$ and $a_{ii} \neq 0$ for all i,

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

1.3   Convergence

Let $\mathbf{x}^*$ be the true solution, i.e., $A\mathbf{x}^* = \mathbf{b}$. Does $\mathbf{x}^{(k)} \mapsto \mathbf{x}^*$ as $k \mapsto \infty$? In other words, does $\mathbf{x}^{(k)} - \mathbf{x}^* \mapsto 0$ as $k \mapsto \infty$?

Let $e^{(k)} = \mathbf{x}^* - \mathbf{x}^{(k)}$. We want to show that $e^{(k)} \mapsto \infty$. So,

$$\begin{aligned} e^{(k+1)} &= \mathbf{x}^* - \mathbf{x}^{(k+1)} \\ &= \mathbf{x}^* - (D^{-1}(\mathbf{b} - (A - D)\mathbf{x}^{(k)})) \\ &= \mathbf{x}^* - (D^{-1}(\mathbf{b} - A\mathbf{x}^{(k)} + D\mathbf{x}^{(k)})) \\ &= \mathbf{x}^* - D^{-1}(\mathbf{b} - A\mathbf{x}^{(k)}) - (D^{-1}D)(\mathbf{x}^{(k)}) \\ &= \mathbf{x}^* - D^{-1}(\mathbf{b} - A\mathbf{x}^{(k)}) - \mathbf{x}^{(k)} \\ &= e^{(k)} - D^{-1}(\mathbf{b} - A\mathbf{x}^{(k)}) && \mathbf{x}^* - \mathbf{x}^{(k)} = e^{(k)} \\ &= e^{(k)} - D^{-1}(\underbrace{\mathbf{b} - A\mathbf{x}^*}_{\mathbf{0}} + \underbrace{A\mathbf{x}^* - A\mathbf{x}^{(k)}}_{Ae^{(k)}}) \\ &= e^{(k)} - D^{-1}Ae^{(k)} \\ &= (I - D^{-1}A)e^{(k)} \end{aligned}$$

Therefore,

$$\begin{aligned} e^{(k+1)} &= (I - D^{-1}A)e^{(k)} \\ &= (I - D^{-1}A)^2 e^{(k-1)} \\ &= (I - D^{-1}A)^3 e^{(k-2)} \\ &= \dots \\ &= (I - D^{-1}A)^{k+1} e^{(0)}, \end{aligned}$$

where $e^{(0)} = \mathbf{x}^* - \mathbf{x}^{(0)}$ and $\mathbf{x}^{(0)}$ is the initial guess we described above. In any case, $e^{(k+1)} \mapsto 0$ as $k \mapsto \infty$ only if $I - D^{-1}A$ is "small."

1.4　Result

If all eigenvalues λ of $B = I - D^{-1}A$ satisfy $|\lambda| < 1$, then the Jacobi method converges to $\mathbf{x}^*$ for every choice of $\mathbf{x}^{(0)}$. Put it differently, if the Jacobi method converges to $\mathbf{x}^*$ for all $\mathbf{x}^{(0)}$, then all eigenvalues λ of B satisfy $|\lambda| < 1$.

Why? Recall from the power method

$$e^{(0)} = \sum_{i=1}^{n} c_i v_i,$$

where the v_i's are the eigenvectors of B. Then,

$$B^k e^{(0)} = \sum_{i=1}^{n} c_i \lambda_i^k v_i = \sum_{i=1}^{n} \lambda_i^k (c_i v_i) \mapsto 0.$$

Note that if all $|\lambda_i| < 1$, then $\lambda_i^k \mapsto 0$ if $k \mapsto \infty$.