1 February 23rd, 2021

1.1 Jacobi for 2D Discrete Laplacian

For the 2D Laplacian, we have:

$$\begin{cases} -u_{xx} - u_{yy} = f & (x,u) \in \Omega = (0,1)^2 \\ u(x,y) = 0 & \text{on } \partial\Omega \end{cases}$$

By central difference:

$$A_2 x = b, x \in \mathbb{R}^N \quad A_2 \in \mathbb{R}^{N \times N}$$

with $N = n^2$ and:

$$A_2 = A \otimes I + I \otimes A$$

, where A is the 1D discrete Laplacian.

Definition 1.1 (**Kronecker Product (Tensor Product)**).

$$B \otimes C = \begin{bmatrix} b_{11}C & b_{12}C & \dots & b_{1q}C \\ \vdots & & & \vdots \\ b_{p1}C & b_{p2}C & \dots & b_{pq}C \end{bmatrix}$$

Theorem 1.2

$$(A \otimes B)(C \otimes D) = (AC) \otimes (CD)$$

Lemma 1.3
The eigenvalues of A_2 are $\lambda_i + \lambda_j$, where λ_i, λ_j are eigenvalues of A and $1 \le i, j \le n$.

Proof. Let (λ_i, u_i) be eigenpairs of A. Then:

$$A_2(u_i \otimes u_j) = (A \otimes I)(u_i \otimes u_j) + (I \otimes A)(u_i \otimes u_j)$$

$$\begin{aligned} A_2(u_i \otimes u_j) &= (A \otimes I)(u_i \otimes u_j) + (I \otimes A)(u_i \otimes u_j) \\ &= Au_i \otimes u_j + u_j \otimes Au_j \\ &= \lambda_i u_i \otimes u_j + u_i \otimes (\lambda_j u_j) \\ &= (\lambda_i + \lambda_j)(u_i \otimes u_j). \end{aligned}$$

□

Thus:

$$\begin{aligned} G_2 &= I - D_2^{-1} A_2 \\ &= I - \frac{1}{4} A_2. \end{aligned}$$

Meaning that the eigenvalues of G_2 are:

$$\begin{aligned} 1 - \frac{1}{2}(\lambda_i + \lambda_j) &= 1 - \frac{1}{4}\left(4 - 2\cos\frac{i\pi}{n+1} - 2\cos\frac{j\pi}{n+1}\right) \\ &= \frac{1}{2}\left(\cos\frac{i\pi}{n+1} + \cos\frac{j\pi}{n+1}\right). \end{aligned}$$

Thus:

$$\rho(G_2) = \max_{1\le i,j\le n}\left|\frac{1}{2}\left(\cos\frac{i\pi}{n+1} + \cos\frac{j\pi}{n+1}\right)\right| = \cos\frac{\pi}{n+1} < 1$$

Because G_2 is symmetric, $\|G_2\|_2 = \rho(G_2)$:

$$\|x_k - x_*\|_2 \le \rho(G_2)\cdot\|x_{k-1} - x_*\|_2$$

Similar to before, we have:

$$1 - O\left(\frac{1}{n^2}\right) = 1 - O\left(\frac{1}{N}\right)$$

This gives us $\alpha = 1$, meaning that:

- number of iterations needed: $O(N\log\tilde{\epsilon}^{-1})$
- number of FLOPs needed per iterations: $O(N)$, which is the number of non-zero entries

Thus the total computation cost is $O(N^2\cdot\log\tilde{\epsilon}^{-1})$. This is the same order as Gaussian Elimination, since $\tilde{\epsilon}$ is usually a constant.

Remark 1.4 — More examples of Jacobi Iteration include the strictly/irreducibly diagonally dominant matrix, which have been covered in MAT5311.

1.2 Jacobi for SPD Matrices

Theorem 1.5

Let $A \in \mathbb{R}^{n\times n}$ be SPD. Then Jacobi converges to x_* for any x_0 if and only if $2D - A$ is SPD too.

Proof. Recall that Jacobi converges to x_* for any x_0 if and only if $\rho(G) < 1$.

- Assume Jacobi converges, then:

$$\rho(I - D^{-1}A) < 1 \iff \rho(I - D^{-\frac{1}{w}}AD^{-\frac{1}{2}}) < 1$$

because $D^{\frac{1}{2}}(I - D^{-1}A)D^{-\frac{1}{2}}$ is similar, thus meaning they share the same eigenvalue and thus spectral radius. Let λ be an eigenvalue of $I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$. Then $|\lambda| < 1$ and $1 + \lambda$ is an eigenvalue of $2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$.

Since A is symmetric, λ is real, meaning that $1 + \lambda$ is positive, thus meaning $2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ is SPD.

Consider:

$$D^{\frac{1}{2}}(2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}} = 2D - A$$

which is also SPD, since they are similar.

- The reverse is very similar, just in reverse. Assume $2D - A$ is SPD. We have:

$$\begin{aligned} 2D - A \text{ is SPD} &\implies 2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \text{ is SPD} \\ &\implies 1 + \lambda, \text{ where } \lambda \text{ is an eigenvalue of } I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}.. \end{aligned}$$

We also have:

$$\begin{aligned} A \text{ is SPD} &\implies D^{-\frac{1}{2}}AD^{-\frac{1}{2}} = I - (I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}) \text{ is SPD} \\ &\implies 1 - \lambda > 0 \\ &\implies \lambda > -1. \end{aligned}$$

where λ is an eigenvalue of $I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$. As such, we have:

$$-1 < \lambda < 1 \implies |\lambda| < 1 \implies \rho(I - D^{-1}A) < 1$$

□

Example 1.6
Consider the 1D Laplacian:

$$\begin{cases} A \text{ is SPD} \\ 2D - A = 4I - A \text{ is SPD} \end{cases} \implies \text{Jacobi converges}$$

1.3 Lower Bound of Jacobi Convergence Rate

We have:

$$\frac{\|x_k - x_*\|}{\|x_{k-1} - x_*\|} \leq \rho(G) + \epsilon$$

for an arbitrarily small ϵ, or $\epsilon = 0$ if G is symmetric. This is a worse-case, i.e. an upper bound. However, this factor is asymptotically optimal, meaning:

$$\lim_{k\to\infty} \frac{\|x_k - x_*\|}{\|x_{k-1} - x_*\|} \geq \rho(G)$$

As such, convergence factor ρ is **tight**.

Let us demonstrate this when G is symmetric first.

Remark 1.7 — For nonsymmetric matrices, this is also true, we will prove later.

$|\lambda_1| > |\lambda_2|$ where λ_1, λ_2 are the largest and 2nd largest eigenvalue of G in absolute value.

Let $G = U \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} U^T$ be the eigenvalue decomposition of G, where $U = \begin{bmatrix} u_1 & u_2 & \dots & u_2 \end{bmatrix}$ is unitary.

Let $x_k - x_* = z_k$. Then:

$$z_k = G^k z_0 = \left(U \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} U^T \right) z_0 = U \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} U^T z_0.$$

Denote $U^T z_0 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$. Thus we have:

$$\begin{aligned} \|z_k\|_2 &= \| \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \| \\ &= \left(\sum_{i=1}^n \lambda_i^{2k} \alpha_i^2 \right)^{1/2} \\ &= |\lambda_1|^k \left(\sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_1} \right)^{2k} \alpha_i^2 \right)^{1/2}. \end{aligned}$$

Thus, we have:

- $\lim_{k\to\infty} \frac{\|z_k\|_2}{\|z_{k-1}\|_2} = \rho(G)$

  Proof.

$$\begin{aligned} \lim_{k\to\infty} \frac{\|z_k\|_2}{\|z_{k-1}\|_2} &= |\lambda_1| \lim_{k\to\infty} \frac{\left(\sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_1} \right)^{2k} \alpha_i^2 \right)^{1/2}}{\left(\sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_1} \right)^{2(k-1)} \alpha_i^2 \right)^{1/2}} \\ &= |\lambda_1| \lim_{k\to\infty} \frac{\alpha_1^2 + \left(\frac{\lambda_2}{\lambda_1} \right)^{2k} \alpha_2^2 + \dots}{\alpha_1^2 + \left(\frac{\lambda_2}{\lambda_1} \right)^{2(k-1)} \alpha_2^2 + \dots} \\ &= |\lambda_1| \\ &= \rho(G). \end{aligned}$$

  □

- $|\langle z_k, u_j \rangle| = |\lambda_j^k \alpha_j| = |\lambda_j|^k |\langle z_0, u_j \rangle|$, meaning that the convergence speed of error projected onto u_j is $|\lambda_j|$.

Remark 1.8 — This means that different error components have different convergence speed.

- The error converges to the same direction as u_1.

Proof.

$$\begin{aligned}
\lim_{k\to\infty} \frac{|\langle z_k, u_j\rangle|}{\|z_k\|_2} &= \lim_{k\to\infty} \frac{|\lambda_j|^k |\langle z_0, u_j\rangle|}{|\lambda_1|^k \left(\sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_1}\right)^{2k} \alpha_i^2\right)^{1/2}} \\
&= \lim_{k\to\infty} \frac{|\lambda_j|^k}{|\lambda_1|^k} \cdot \frac{|\alpha_1|}{|\alpha_1|} \\
&= \lim_{k\to\infty} \left|\frac{\lambda_j}{\lambda_1}\right| \\
&= \begin{cases} 1 & j = 1 \\ 0 & j \neq 0 \end{cases}.
\end{aligned}$$

□

Example 1.9

Let's consider the 1D discrete Laplacian:

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix}, \quad G = I - \frac{1}{2}A$$

Recall that the eigenvalues of A are :

$$\lambda_j(A) = 2\left(1 - \cos\frac{j\pi}{n+1}\right), \quad j = 1, 2 \ldots, n$$

meaning that the eigenvalues of G aer:

$$\lambda_j(G) = 1 - \frac{1}{2}\lambda_j(A) = \cos\frac{j\pi}{n+1}, \quad j = 1, 2 \ldots, n$$

Let u_j be the eigenvectors of A (and of G, since it is a shift of A). Let the error be

$$z_k = x_k - x_* = \sum_{j=1}^{n} \alpha_j^{(k)} u_j$$

Then the convergence speed of $\alpha_j^{(k)}$ is $\lambda_j(G)$. As such, we have:

- If j is close to 1 or n, then $|\lambda_j(G)|$ is close to 1, meaning it has a slower convergence.
- If j is close $\frac{n}{2}$, then $|\lambda_j(G)|$ is close to 0, meaning it has a faster convergence.

Remark 1.10 — To improve Jacobi, in the Multigrid, we use different grids. The main motivation is that the convergence speed of different components depends on the eigenvalues. For components that are difficult to decrease, we project them onto a coarse grid to reduce the error.