FM 5012
4th class

Classic Iterative Methods. *

① Description of the methods.

The classic iterative methods to solve the matrix equation $Ax = b$, use a splitting of the matrix A, namely,

$$A = N - P$$

and use it to generate a sequence $\{x^n\}_{n\geq 0}$ that should converge to the solution x. the sequence is generated as follows:

$$(I) \begin{cases} \text{① Pick } x^0 = x_0. \\ \text{② If } r^k = Ax^k - b = 0 \text{ stop, otherwise} \\ \quad \text{compute } x^{k+1} \text{ by solving} \\ \quad N x^{k+1} = P x^k + b. \end{cases}$$

Note that if $e^k = x - x^k$, we get that the error equation is

$$N e^{k+1} = P e^k \qquad k \geq 0.$$

The matrix $M = N^{-1}P$ is called the iteration matrix. the convergence of the method (I) depends only on properties of M.

* From: Introd. à l'analyse numérique et à l'optimization, P.G. Ciarlet

The classic iterative methods are described in the table below, we use the fact that D is the diagonal part of A, U is the upper part and L the lower part

Name	iteration matrix
Jacobi	$-D^{-1}(A-D)$
Gauss-Seidel	$-(D+L)^{-1}(U)$
SOR	$(D+\omega L)^{-1}(D-\omega(A-L)) = M_{SOR}(L)$
SSOR	$M_{SOR}(U)\, M_{SOR}(L)$

Note that:

① $M_{SOR}(L)\big|_{\omega=1} = M_{GS}$.

② the SOR can be rewritten as:

$$\begin{cases} D\,\tilde{x}^{k+1} = -L\,x^{k+1} - U\,x + b \\ x^{k+1} = (1-\omega)\,x^k + \omega\,\tilde{x}^{k+1} \end{cases}$$

③ M_{SSOR} is symmetric if A is symmetric.
M_{SOR} is not symmetric even if A is.

② Strictly diagonal dominant matrices.

A matrix A is said to be strictly diagonal dominant if

$$\rho = \max_{1 \le i \le N} \sum_{\substack{j=1 \\ j \neq i}} \frac{|a_{ij}|}{|a_{ii}|} < 1.$$

For these matrices, the error of both the Jacobi and the Gauss-Seidel methods satisfy

$$\|e^n\|_{l^\infty} \le \rho^n \|e^0\|_{l^\infty}.$$

③ Symmetric and positive definite matrices.

For these matrices, we have that

(i) the Jacobi method converges if $a_{ij} \le 0$ for $i \neq j$.

(ii) the SOR method converges if $\omega \in (0,2)$. It diverges if $\omega \in (-\infty, 0) \cup (2, \infty)$.

to show this result, we prove the following theorem.

Theorem. Let A be symmetric and positive definite. Suppose that

$$Q = N + N^T - A$$

is positive definite. Then $\rho(M) < 1$.

Proof. Let λ be an eigenvalue of M and let u be its eigenvector. Then

$$\begin{aligned} & Mu = \lambda u \\ \Rightarrow \quad & (I - N^{-1}A)u = \lambda u \\ \Rightarrow \quad & Au = (1-\lambda) Nu \end{aligned}$$

Since A is positive definite $Au \neq 0$ and hence $\lambda \neq 1$.

Next, let us exploit the fact that Q is positive definite:

$$\begin{aligned} 0 < \overline{u}^T Q u &= \overline{u}^T (N + N^T - A) u \\ &= \overline{u}^T N u + \overline{u}^T N^T u - \overline{u}^T A u \\ &= \frac{\overline{u}^T A u}{1-\lambda} + \frac{\overline{u}^T A u}{1-\overline{\lambda}} - \overline{u}^T A u \\ &= \frac{1 - \lambda\overline{\lambda}}{(1-\lambda)(1-\overline{\lambda})} \qquad \Rightarrow \quad |\lambda| < 1. \\ & \qquad\qquad\qquad\qquad\quad \Rightarrow \quad \rho(M) < 1. \end{aligned}$$

■

Let us prove the claims about the convergence of the Jacobi and SOR methods. We assume that A is real.

Let us begin with the Jacobi method. In this case $N = D$ and

$$Q = N + N^T - A = 2D - A = |A|.$$

Since A is positive definite, so is $|A|$. Hence Q is positive definite.

Let us now consider the SOR method. In this case, we have

$$N = \frac{1}{\omega}(D + \omega L).$$

Hence

$$\begin{aligned} Q &= \frac{1}{\omega}(D + \omega L + D + \omega U) - A \\ &= \left(\frac{2}{\omega} - 1\right) D, \end{aligned}$$

which is positive definite if $0 < \omega < 2$, as claimed. Now, note that

$$\begin{aligned} \det M_{SOR} &= \det\left((D + \omega L)^{-1}(D - \omega(D + U))\right) \\ &= \det\left((I + \omega D^{-1}L)^{-1}(Id - \omega(Id + D^{-1}U)\right) \\ &= (1 - \omega)^N \end{aligned}$$

and so, $\rho(M_{SOR}) \geq |1-\omega|$. this implies that the SOR method diverges if $\omega \notin [0,2]$.

Note that

$$\rho(M_{SSOR}) \leq \rho^2(M_{SOR}).$$

and so, the SSOR method converges for $\omega \in (0,2)$. An argument similar to that used for the SOR method shows that for $\omega \notin [0,2]$, the SSOR method diverges.

④ Symmetric, positive definite matrices that are block-tridiagonal.

For these matrices, we have the following result:

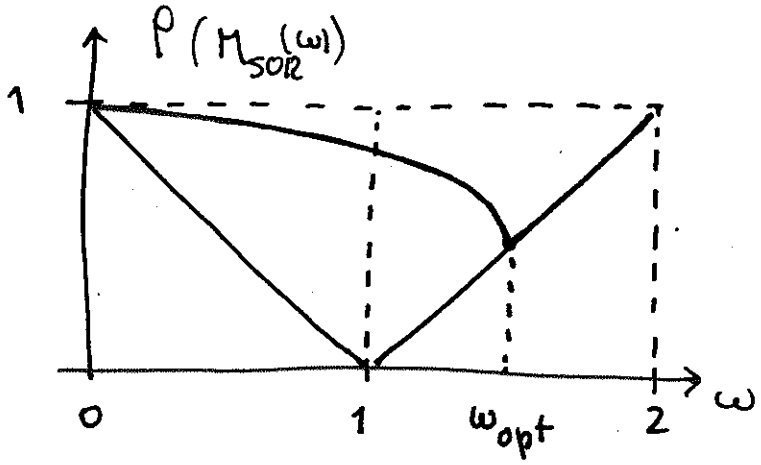
the Jacobi method, the Gauss-Seided method and the SOR with $\omega \in (0,2)$ converge. Moreover, we have:

$$\begin{aligned}\rho(M_{SOR}(\omega_{opt})) &= \inf_{\omega \in \mathbb{R}} \rho(M_{SOR}(\omega)) \\ &= \omega_{opt} - 1 \\ &< \rho(M_{GS}) \\ &= \rho^2(M_J) \quad < \rho(M_J),\end{aligned}$$

where

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho^2(M_J)}} .$$

Note that



to prove this result, we use a fundamental fact for block-tridiagonal matrices. Consider the matrix

$$A(\mu) = \begin{bmatrix} D_1 & \tilde{\mu}^{-1} U_1 & 0 & & \\ \mu L_2 & D_2 & \tilde{\mu}^{-1} U_2 & & \\ 0 & \mu L_3 & D_3 & \ddots & \\ & & \ddots & \ddots & \tilde{\mu}^{-1} U_{N-1} \\ & & & \mu U_N & D_N \end{bmatrix}$$

$$= D + \mu L + \tilde{\mu}^{-1} U .$$

then

$$\det A(\mu) = \det A(0)$$

or, equivalently,

$$\det\left(D+\mu L+\mu^{-1}U\right)=\det(D+L+U). \tag{*}$$

To prove this result, it is enough to realize that

$$A(\mu)=\Phi(\mu)\,A(1)\,\Phi(\mu)^{-1}$$

where

$$\Phi(\mu)=\begin{bmatrix}\mu I_1 & & 0\\ & \mu^2 I_2 & \\ & & \ddots \\ 0 & & \mu^N I_N\end{bmatrix}.$$

Now, to prove the convergence result, we proceed as follows. First, let us consider the following polynomial

$$p_J(\lambda)=\det\left(-D^{-1}(U+L)-\lambda Id\right).$$

By definition, the zeroes of p_J are the eigenvalues of the iteration matrix for the Jacobi method. Then

$$p_J(\lambda)=\det\left(-D^{-1}\right)\cdot\det\left(U+L+\lambda D\right).$$

Now, let us consider the SOR method:

$$p_{SOR}(\lambda) = \det\left((D+\omega L)^{-1}((1-\omega)D-\omega U) - \lambda\, Id\right)$$
$$= \det\left(\left(\frac{D}{\omega}+L\right)^{-1}\left(\frac{1-\omega}{\omega}D-U\right) - \lambda\, Id\right)$$

$$= \det\left(\frac{D}{\omega}+L\right)^{-1} \cdot \det\left(\frac{1-\omega}{\omega}D-U-\frac{\lambda}{\omega}D-\lambda L\right)$$

$$= \det\left(\frac{D}{\omega}+L\right)^{-1} \cdot \det\left(\frac{1-\lambda-\omega}{\omega}D-U-\lambda L\right)$$

$$= \det\left(\frac{D}{\omega}+L\right)^{-1} \cdot (-\sqrt{\lambda})^{N} \cdot \det\left(\frac{\lambda+\omega-1}{\omega\sqrt{\lambda}}D+\frac{1}{\sqrt{\lambda}}U+\sqrt{\lambda}L\right)$$

by the property (*),

$$p_{SOR}(\lambda) = \det\left(\frac{D}{\omega}+L\right)^{-1}(-\sqrt{\lambda})^{N} \cdot \det\left(\frac{\lambda+\omega-1}{\omega\sqrt{\lambda}}D+U+L\right),$$

and hence

$$p_{SOR}(\lambda) = C(\lambda)\; p_J\left(\frac{\lambda+\omega-1}{\omega\sqrt{\lambda}}\right),$$
$$C(\lambda) = \det\left(\frac{D}{\omega}+L\right)^{-1} \cdot \det D^{-1} \cdot \lambda^{N/2}.$$

this means that

$$\alpha \in \text{spectrum of } H_J$$
$$\Leftrightarrow \quad \lambda:\ \frac{\lambda+\omega-1}{\omega\sqrt{\lambda}} = \alpha \quad \in \text{spectrum of } H_{SOR}(\omega)!$$

If $\omega=1$, $\sqrt{\lambda}=\alpha$ and so

$$\rho(M_{SOR(\omega=1)}) = \rho(M_{GS}) = \rho^2(M_J).$$

To prove the remainder of the result, let us show that $\alpha \in$ spectrum of M_J is real. Indeed, if we set that

$$-\bar{D}^{-1}(U+L)\,v = \alpha v,$$

then

$$\begin{aligned} 0 &= (U+L+D\alpha)\,v \\ &= (A+(\alpha-1)D)\,v \end{aligned}$$

and so

$$\bar{v}^T A v^T = (1-\alpha)\,\bar{v}^T D v,$$

which proves the claim since A is symmetric and positive definite. Moreover, $\alpha \in [0,1)$.

As a consequence, we must study the number

$$\rho(M_{SOR}^{(\omega)}) = \max_{\alpha \in \text{spectrum } M_J} \left\{ |\lambda| : \frac{\lambda+\omega-1}{\sqrt{\lambda}\omega} = \alpha \right\}.$$

assuming that spectrum $M_J \subset [0,1)$.

the theorem follows from such study.

⑤ the model problem.

Let Ω be the unit square and let U_{ij} denote the classic finite difference approximation to the exact solution of

$$(P) \qquad -\Delta u = f \text{ in } \Omega; \quad u = 0 \text{ on } \partial\Omega,$$

given by

$$(P_h) \begin{cases} -\frac{1}{h^2}\left(U_{i+1,j} + U_{i-1,j} + U_{i,j+1} + U_{i,j-1} - 4U_{ij}\right) = f_{ij} \\ i,j = 1, \ldots, N-1 \\ U_{ij} = 0 \quad \text{if } i = 0,\ j = 0,\ i = N \text{ or } j = N. \end{cases}$$

It is not difficult to show that if A denotes the matrix associated to (P_h), it is positive definite and symmetric. Moreover,

$$AU = \lambda U$$

$$\Leftrightarrow \qquad U = U^{nm}, \quad \lambda = \lambda^{nm}$$

$$U_{ij}^{nm} = (\operatorname{sen} n\pi x_i)(\operatorname{sen} m\pi y_j)$$

$$\lambda^{mn} = \frac{4}{h^2}\left(\operatorname{sen}^2 \frac{m\pi h}{2} + \operatorname{sen}^2 \frac{n\pi h}{2}\right)$$

where $x_i = ih$, $y_j = jh$.

In this case, we can write

$$e^k = \sum_{m,n=1}^{N-1} c_{mn}^k \, v^{mn},$$

where $e^k = x - x^k$ and x^k is the k-th iterate of the Jacobi method. Then, since

$$\begin{aligned} e^{k+1} &= -D^{-1}(U+L)\, e^k \\ &= D^{-1}(D-A)\, e^k \\ &= (Id - D^{-1}A)\, e^k \end{aligned}$$

we get that

$$c_{mn}^{k+1} = \left(1 - \frac{h^2}{4}\lambda^{mn}\right) e_{mn}^k .$$

This implies that

$$c_{mn}^{k+1} = \left(1 - \sin^2\frac{mh\pi}{2} - \sin^2\frac{nh\pi}{2}\right) c_{mn}^k .$$

If $m = n = N/2$,

$$\begin{aligned} c_{mn}^{k+1} &= \left(1 - 2\sin^2\frac{Nh}{4}\pi\right) c_{mn}^k \\ &= 0, \end{aligned}$$

and this implies that the components of the error associated with v^{mn}, $m,n \approx N/2$ are damped extremely fast.

If $m=n=1$,

$$c_{mn}^{k+1} = \left(1-2\sin^2\frac{\pi h}{2}\right) c_{mn}^{k}$$
$$= \left(1-\frac{\pi^2 h^2}{2}+O(h^4)\right) c_{mn}^{k},$$

and if $m=n=N-1$,

$$c_{mn}^{k+1} = \left(1-2\sin^2\frac{(N-1)\pi h}{2}\right) c_{mn}^{k}$$
$$= -\left(1-2\sin^2\frac{\pi h}{2}\right) c_{mn}^{k}$$
$$= \left(-1+\frac{\pi^2 h^2}{2}+O(h^4)\right) c_{mn}^{k},$$

which means that for $m\tilde{=}n\tilde{=}1$ and $m \text{ or } n \neq N-1$, the corresponding frequencies are barely damped.