Dr A. Hansen

Mathematical Tripos Part II: Michaelmas Term 2020

Numerical Analysis – Lecture 18

Observation 4.17 (Attenuation for different frequencies) The speed of convergence of some iterative methods (Jacobi with relaxation, Gauss–Seidel, etc.) can be increased drastically within the context of solving linear equations that originate in the discretization of PDEs. Herewith we analyse (with a great deal of hand-waving) the system $A\boldsymbol{u} = \boldsymbol{b}$ originated from the 5-point formula for the Poisson equation on an $m \times m$ square grid Ω_h, being solved by the *damped* Jacobi iteration.

Method 4.18 (The damped Jacobi iteration) This is the Jacobi method with a relaxation parameter ω (see Lect. 17):

$$
\begin{aligned}
\widehat{\boldsymbol{u}}^{(\nu+1)} &= -D^{-1}(A-D)\boldsymbol{u}^{(\nu)} + D^{-1}\boldsymbol{b} &&= (I - D^{-1}A)\boldsymbol{u}^{(\nu)} + D^{-1}\boldsymbol{b} \\
\boldsymbol{u}^{(\nu+1)} &= \omega\widehat{\boldsymbol{u}}^{(\nu+1)} + (1-\omega)\boldsymbol{u}^{(\nu)} &&= (I - \omega D^{-1}A)\boldsymbol{u}^{(\nu)} + \omega D^{-1}\boldsymbol{b}\,.
\end{aligned}
$$

The error decay is expressed in terms of the iteration matrix H_ω:

$$
\boldsymbol{e}^{(\nu)} = [H_\omega]^\nu \boldsymbol{e}^{(0)}, \qquad H_\omega = I - \omega D^{-1}A = I + \tfrac{1}{4}\omega A\,,
$$

and it follows from the results of Lecture 2 that the eigenvectors and the eigenvalues of H_ω are

$$
\boldsymbol{w}^{k,\ell} = (\sin ix \sin jy), \qquad \lambda_{k,\ell}(\omega) = 1 - \omega\left(\sin^2 \tfrac{x}{2} + \sin^2 \tfrac{y}{2}\right), \qquad x = \tfrac{k\pi}{m+1}, \quad y = \tfrac{\ell\pi}{m+1}\,.
$$

We see that $\rho(H_\omega) < 1$ for any ω in $(0,1]$, guaranteeing convergence, although a very slow one. In particular, for the "pure" Jacobi iteration (with $\omega = 1$) we have $\rho(H_J) = 1 - 2\sin^2 \tfrac{\pi}{2(m+1)} \approx 1 - \tfrac{\pi^2}{2m^2}$, and for $\omega < 1$ the spectral radius is even closer to 1.

However, expanding the error with respect to the (orthogonal) eigenvectors we obtain

$$
\boldsymbol{e}^{(\nu)} = \sum\nolimits_{k,\ell} a_{k,\ell}^{(\nu)} \boldsymbol{w}^{k,\ell}, \qquad \boldsymbol{e}^{(\nu)} = [H_\omega]^\nu \boldsymbol{e}^{(0)} \;\Rightarrow\; |a_{k,\ell}^{(\nu)}| = |\lambda_{k,\ell}(\omega)|^\nu |a_{k,\ell}^{(0)}|\,,
$$

i.e. the components of $\boldsymbol{e}^{(\nu)}$ (with respect to the basis of eigenvectors) decay at a different rate for different frequences (k, ℓ). To this end, we define

$$
\begin{aligned}
&\Omega_h\text{-low frequencies (LF):} &&\boldsymbol{w}^{(k,\ell)} = (\sin \pi k\phi \sin \pi\ell\psi)\big|_{\Omega_h} &&\text{with both } k \text{ and } \ell \text{ from } [1, \tfrac{m+1}{2})\text{,} \\
&\Omega_h\text{-high frequencies (HF):} &&\boldsymbol{w}^{(k,\ell)} = (\sin \pi k\phi \sin \pi\ell\psi)\big|_{\Omega_h} &&\text{with either } k \text{ or } \ell \text{ from } [\tfrac{m+1}{2}, m]\text{.}
\end{aligned}
$$

Let us determine the least factor $\mu(\omega)$ by which the amplitudes of HF components are damped per iteration. We have

$$
\begin{aligned}
\mu(\omega) &= \max\left\{|\lambda_{k,\ell}(\omega)| : \quad \tfrac{m+1}{2} \le k \le m, \quad 1 \le \ell \le m\right\} \\
&= \max\left\{|1 - \omega(\sin^2 \tfrac{x}{2} + \sin^2 \tfrac{y}{2})| : \quad \tfrac{\pi}{2} \le x \le \pi, \quad 0 \le y \le \pi\right\} \\
&= \max\left\{|1 - \tfrac{1}{2}\omega|, |1 - 2\omega|\right\},
\end{aligned}
$$

and it is seen that the optimal factor μ_* is attained when $1 - \tfrac{1}{2}\omega = -(1 - 2\omega)$, i.e. for $\omega_* = \tfrac{4}{5}$, and its value is $\mu_* = \tfrac{3}{5}$. Therefore, for the coefficients at the HF components of $\boldsymbol{e}^{(\nu)}$ we obtain

$$
|a_{k,\ell}^{(\nu)}| \le |\mu_*|^\nu |a_{k,\ell}^{(0)}| = \left(\tfrac{3}{5}\right)^\nu |a_{k,\ell}^{(0)}| \ll |a_{k,\ell}^{(0)}|\,,
$$

i.e. the damped Jacobi method converges fast for high frequencies.

For the remaining Ω_h-low frequences we notice that

$$
k, \ell \in [1, \tfrac{m+1}{2}) = [1, \tfrac{1}{2h}) \;\Rightarrow\; \underbrace{\{\boldsymbol{w}_h^{(k,\ell)} = (\sin \pi k\phi \sin \pi\ell\psi)\big|_{\Omega_h}\}}_{\Omega_h\text{-low frequences}} \approx \underbrace{\{\boldsymbol{w}_{2h}^{(k,\ell)} = (\sin \pi k\phi \sin \pi\ell\psi)\big|_{\Omega_{2h}}\}}_{\Omega_{2h}\text{-high frequences}}
$$

Algorithm 4.19 (The multigrid method) The idea of the multigrid method is that, although the global error may decrease slowly by iteration, its components with high frequencies relative to Ω_h are suppressed (or smoothed) very quickly, and that dealing with the remaining components (with low frequences relative to Ω_h) we can move to the coarse grid Ω_{2h}, where these components (in part) would be of high frequencies, and thus they can be smoothed in a similar way. Therefore, we cover the square domain by a range of nested grids, of increasing coarseness, say,

$$\Omega_h \subset \Omega_{2h} \subset \Omega_{4h} \subset \cdots \subset \Omega_{2^m h}.$$

At every Ω_{h_i}, the iterations (damped Jacobi, or Gauss-Seidel) remove the high frequencies relative to this grid, and we move to Ω_{2h_i}. On the coarsest grid, where the number of variables is small, we can afford to solve the equations with a direct method, by Cholesky, say.

A *multigrid sweep* starts at the finest grid, travels to the coarsest (where we apply a direct solver) and back to the finest:

1) Each *coarsening* stage involves few (ν) iterations, then computing the residual $\boldsymbol{r}_h^{(\nu)} = \boldsymbol{b}_h - A_h \boldsymbol{u}_h^{(\nu)}$ (where h is the size of the grid Ω_h) and restricting it to the coarser grid Ω_{2h} via "restriction" mapping $R : \mathbb{R}^N \to \mathbb{R}^{N/4}$ as $\boldsymbol{r}_{2h}^{(0)} = R \boldsymbol{r}_h^{(\nu)}$:

$$
\begin{aligned}
\boldsymbol{r}_h^{(\nu)} &= \boldsymbol{b}_h - A_h \boldsymbol{u}_h^{(\nu)} = A_h \boldsymbol{e}_h^{(\nu)} = \sum_{1 \le k < m+1} \sum_{1 \le \ell < m+1} a_{k,\ell}^{(\nu)} \boldsymbol{w}_h^{k,\ell} \\
&= \sum_{1 \le k < \frac{1}{h}} \sum_{1 \le \ell < \frac{1}{h}} a_{k,\ell}^{(\nu)} \sin \pi k \phi \sin \pi \ell \psi \Big|_{\Omega_h} \\
&= \underbrace{\sum_{1 \le k < \frac{1}{2h}} \sum_{1 \le \ell < \frac{1}{2h}} a_{k,\ell}^{(\nu)} \sin \pi k \phi \sin \pi \ell \psi \Big|_{\Omega_h}}_{\text{transferred (restricted) to } \Omega_{2h} \text{ as}} + \underbrace{\left(\sum_{\frac{1}{2h} \le k < \frac{1}{h}} \sum_{1 \le \ell < \frac{1}{h}} + \sum_{1 \le k < \frac{1}{h}} \sum_{\frac{1}{2h} \le \ell < \frac{1}{h}} \right) a_{k,\ell}^{(\nu)} \sin \pi k \phi \sin \pi \ell \psi \Big|_{\Omega_h}}_{\text{suppressed to } \boldsymbol{\epsilon}_h^{(\nu)} \text{ and transferred to } \Omega_{2h} \text{ as}}
\end{aligned}
$$

$$
\underbrace{\sum_{1 \le k < \frac{1}{2h}} \sum_{1 \le \ell < \frac{1}{2h}} a_{k,\ell}^{(\nu)} \sin \pi k \phi \sin \pi \ell \psi \Big|_{\Omega_{2h}}}_{} + \qquad \boldsymbol{\epsilon}_{2h}^{(\nu)} \qquad =: \boldsymbol{r}_{2h}^{(0)}
$$

A typical restriction operator R combines nine "fine" values around the "coarse" one according to the rule:
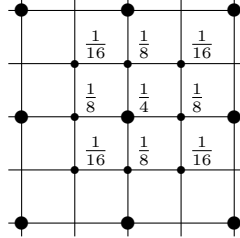
At the coarse grid Ω_{2h}, we are solving *for the residual*, i.e. we iterate μ times for (or apply a direct solver to) the low-dimensional system

$$A_{2h} \boldsymbol{v}_{2h} = 4 \, \boldsymbol{r}_{2h}^{(0)}.$$

The factor 4 on the right is due to the fact that our linear system is of the form $A_h \boldsymbol{u}_h = h^2 \boldsymbol{f}_h$, hence its coarse version is $A_{2h} \boldsymbol{v}_{2h} = (2h)^2 \boldsymbol{g}_{2h}$. Having found approximative (or exact solution) $\boldsymbol{v}_{2h}^{(\mu)}$ we move back to Ω_h.

2) Refinement entails a prolongation via mapping $P : \mathbb{R}^{N/4} \to \mathbb{R}^N$ as

$$\boldsymbol{v}_h^{(\mu)} = P \boldsymbol{v}_{2h}^{(\mu)},$$

e.g. by linear interpolation (the exact opposite of the above procedure), and correction

$$\boldsymbol{u}_h^{\text{new}} = \boldsymbol{u}_h^{(\nu)} + \boldsymbol{v}_h^{(\mu)}.$$

The resulting vector is close to the exact solution,

$$\boldsymbol{u}_h^{(\nu)} + P A_{2h}^{-1} R \boldsymbol{r}_h^{(\nu)} \approx \boldsymbol{u}_h^{(\nu)} + A_h^{-1}(b_h - A_h \boldsymbol{u}_h^{(\nu)}) = \boldsymbol{u}_h^*.$$

It is usual to employ only a moderate number of iterations in each restriction (3–5, say) and prolongation (just 1–2 iterations, to take care of high frequencies that have been reintroduced by prolongation) and to check for convergence only by the end of the sweep. Unless convergence occurs, we embark on another multigrid sweep and so on.