Multigrid

onsdag 1. november 2017 15:37

1. We begin with considering JOR method applied to $Au = b$ (Jacobi over-relaxation)

$\Rightarrow u^{(k+1)} = B u^{(k)} + \alpha P^{-1} b$

where $B = (I - \alpha P^{-1} A)$

JOR $\Rightarrow P = D$, diagonal part of A

→ Suppose that A is a linear finite element discretization of $\nabla^2 u$ on a uniform grid (i.e. coinciding with 5-point finite difference method).

$\Rightarrow$ B has eigenvalues / eigenvectors

$$w^{kl} = \sin \frac{ik\pi}{m+1} \sin \frac{jl\pi}{m+1} \quad (\text{i.e. } w_{kl}(x_i, y_j))$$

$$\lambda_{kl} = 1 - \alpha \left(\sin^2 \frac{k\pi}{2(m+1)} + \sin^2 \frac{l\pi}{2(m+1)} \right)$$

$(k, l = 1, \ldots, m)$.

→ We now expand the error in terms of this eigenbasis, i.e.

$$e^{(n)} = \sum_{k,l} a_{k,l}^{(n)} w^{kl}, \quad \text{for some } a_{k,l}^{(n)}.$$

Now recall that

$$e^{(n)} = B^n e^{(0)}$$

and as the w^{kl} are eigenvectors for B, we have

$$e^{(n)} = \sum \lambda_{kl}^{n} a_{kl}^{(0)} w^{kl}$$

we have

$$e^{(n)} = \sum_{k,l} \delta_{k,l}^{n} \, a_{k,l}^{(0)} \, w^{kl}$$

→ Recall that $\delta_{k,l} = 1 - \alpha \left(\sin^2 \frac{k\pi}{2m+2} + \sin^2 \frac{l\pi}{2m+2} \right)$

we are going to fix α so that the components of 'high frequency', i.e. such that at least one of $k, l > \frac{m}{2}$ are smoothed quickly.

Now,

high frequencies

$$\sup_{k,l} |\delta_{k,l}| , \; k \in [\tfrac{m}{2}, m], \; l \in [1, m]$$

$$= \max \left\{ \left|1 - \frac{\alpha}{2}\right| , |1 - 2\alpha| \right\}$$

$\Rightarrow$ max for $1 - \frac{\alpha}{2} = 1 - 2\alpha \Rightarrow \alpha = \frac{4}{5}$

$\Rightarrow$ for optimal $\alpha = \frac{4}{5}$, have

$$\underbrace{\sup_{k,l}}_{\text{high frequencies}} |\delta_{k,l}| = \frac{3}{5}$$

→ To conclude, high frequencies are attenuated like

$$|a_{k,l}^{(n)}| = \underbrace{\left(\frac{3}{5}\right)^n}_{\text{rapid decay?}} |a_{k,l}^{(0)}|$$

Much faster than $1 - h^2$

---

2. The main idea of multigrid methods is to exploit the fact that the

splitting into high/low frequencies changes with the grid size

2) roughly speaking, we can consider solving m problems

$$A_h u_h = b_h ,$$ for different values of h (associated grids T_h).

the idea is to set up procedures mapping $u_{h_1} \longmapsto u_{h_2}$

(called interpolation/refinement if $h_1 > h_2$

and restriction/coarsening if $h_1 < h_2$)

we then perform a few iterations of, eg, JOR, before moving to a different grid by one of the above strategies

we then run JOR on the new grid etc.

3. Multigrid schemes:

<u>V-cycle:</u>

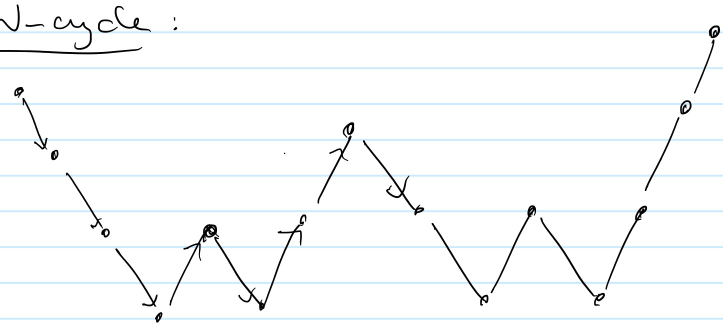
$$u_{h_1}^{(0)} \to u_{h_2}^{(0)} \to u_{h_3}^{(0)} \to u_{h_4}^{(0)} \to u_{h_3}^{(1)} \to u_{h_2}^{(1)} \to u_{h_1}^{(1)}$$

where $h_1 < h_2 < h_3 < h_4$

(i.e. we sweep from finest to coarsest grid and back)

NB: the choice of 4 here is purely illustrative, in general for $h_1 < \dots < h_n$ the pattern is the same

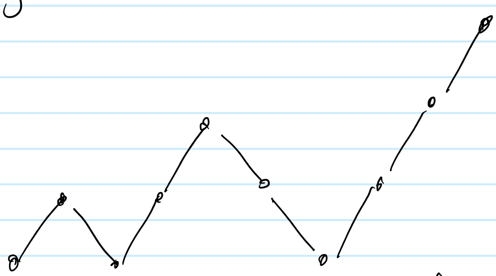
W-cycle:



etc.

(this often performs better in less regular cases, eg where A is not symmetric positive definite)

Full multigrid: here the first round begins from the coarsest level, eg.



$\longrightarrow$ onto V-cycles

(or similar with W-cycles).

---

3. Some convergence results

— First multigrid theorem:

Consider the solution $u \in H^1$ of the equation

$$a(u,v) = f(v), \qquad \forall v \in H^1$$

where a is symmetric, continuous, coercive bilinear form

f is a continuous, linear functional defined by $f(v) = \int_\Omega gv$, some $g \in L_2$.

Assume :

i) We construct $(k+1)$ linear systems

$$A_{h_i} u_i = F_i$$

from $(k+1)$ different grids T_{h_i}

$(h_i = 0, \ldots, k)$

with parameters $h_0 > \ldots > h_k > 0$

: $\frac{h_i}{h_{i+1}} \leq C$ ← some constant.

ii) The discretized systems are constructed by the Galerkin method with basis functions $\varphi_j^{h_i}$.

and moreover, each $\varphi_j^{h_i}$ is a finite linear combination of some $\varphi_{j\ell}^{h_{i+1}}$

(this will be satisfied whenever the finer grid consists of subdivisions of the elements of the coarse grid, for instance)

iii) The exact solution is related to the approximate solutions by an estimate

$$\| u - u^{h_i} \|_{L^2} \leq c \cdot h_i^2 \| g \|_{L^2}$$

(this will be satisfied in the presence of the regularity estimate

$$|u|_{H^2} \leq C \cdot \|g\|_{L^2}$$

which we recall was used for the derivation of the L^2 convergence theorem under the name elliptic regularity).

iv) The eigenvalues of A_{h_i} obey the bound $\lambda(A_{h_i}) \leq C \cdot h_i^{d-2}$

(this is always satisfied for us, it is derived similarly to the condition number estimate).

We now introduce the vector norm

$$\|u\|_{h_i} = h_i^{d/2} |u|$$

d = dim of space.

$|u|$: Euclidean vector norm

(i.e. this is the vector norm reflecting the L^2 norm of associated functions $\sum_k u_k \varphi_k^i(x)$.)

<u>Theorem</u>: under the above assumption, perform JOR on each downward step of a W-cycle. Assume that interpolation is performed by writing $\varphi_j^{h_{i+1}} = \sum \alpha_k \varphi_k^{h_i}$.

(i.e. the natural interpolation)

Then $\forall \varepsilon \in (0,1)$, $\exists m$: if m JOR iterations are performed at each step, the full cycle reduces the error by a factor of ε:

$$\|u_{(1)}^{h_i} - u^{h_i}\|_E \leq \varepsilon \|u_{(0)}^{h_i} - u^{h_i}\|_E.$$

We will now use this to give a convergence result for the full scheme:

suppose in addition that the functional $f(v)$ is such that

$$\| u - u^{h_i} \|_{L^2} \leqslant C . h_i^2$$

Theorem: under the above assumptions, we obtain estimates of the form

$$\| u - u^{h_i} \|_{L^2} \leqslant C(t) . h_i^2$$

where t is the number of multigrid cycles executed.

(NB : in general, the number of iterations m required is related to ε by

$$\varepsilon = \mathcal{O}(m^{-2}) \;).$$

Note that a similar result holds for the error in energy norm, i.e.

$$\| u \|_E = \sqrt{a(u,u)} \, ,$$

but here as well we require an elliptic regularity-type result:

$$| u |_{H^2} \leqslant c \| g \|_{L^2} .$$

- Similar results can be derived for the non-symmetric case, but speed of convergence is typically lower

(e.g. $\varepsilon = \mathcal{O}(m^{-1})$).

4. Practical points: we can find amazing results such as the following

- Theorem: Suppose we solve a linear, elliptic, second order PDE with homogeneous Dirichlet conditions on a convex polygon Ω, by linear finite elements.

(assume the forcing $f \in L^2$ etc.)

Then application of the full multigrid algorithm results in estimates

$$\| u - u^{h_i} \|_{L^2} \leq h_i^2 \| f \|_{L^2}$$

where the number of FLOPS (floating point operations) taken in the solution of the linear equations scales linearly with the number of nodes N_i.

(i.e., we can essentially solve a linear system $A_h u = F$ with complexity $O(N)$, not $O(N^3)$ as usual!)

- Here we assumed use of JOR as the iteration. Alternatives exist. In particular, it is often better to switch to conjugate gradient on the coarser stages of the multigrid cycle.

- The results presented here just touch

the surface of multigrid theory!