Dear Sue and Ridgway:

I'm teaching multigrid again using the chapter from your book as a reference, and I have some observations, which you may or may not find relevant (as the final observation is that the way your book does it is better than the way I've been doing it over the years).

Basically, I've been lazy, and I didn't want to write a separate direct solver.

So the code in the package my students use for projects is:

$MG(k, z_0, g)$:

Presmoothing Step. For $1 \le l \le m_1$, let

$$z_l = z_{l-1} + \frac{1}{\Lambda_k}\big(g - A_k z_{l-1}\big),$$

where $\Lambda_k = \|A_k\|_\infty$ satisfies $\lambda_{\max} \le \Lambda_k \le Ch_k^{-2}$, where $\lambda_{\max}$ is the largest eigenvalue of A_k.

Error Correction Step. If $k > 1$, then let $\bar{g} = I_k^{k-1}\big(g - A_k z_{m1}\big)$ and $q_0 = 0$, and for $1 \le i \le p$, let

$$q_i = MG(k-1, q_{i-1}, \bar{g});$$

then

$$z_{m_1+1} = z_{m_1} + I_{k-1}^k q_p.$$

Otherwise, $z_{m_1+1} = z_{m_1}$.

Postsmoothing Step. For $m_1 + 1 < l \le m_1 + m_2 + 1$, let

$$z_l = z_{l-1} + \frac{1}{\Lambda_k}\big(g - A_k z_{l-1}\big).$$

Finally. Return $z_{m_1+m_2+1}$.

So, at the coarsest level I don't do a direct solve and just smooth $m_1 + m_2$ times and leave it at that.

For full multigrid I do

Let $\hat{u}_0 = 0$. For $k \ge 1$ do

$$\begin{aligned} u_0^k &= \hat{u}_{k-1}, \\ u_l^k &= MG(k, u_{l-1}^k, f_k), \quad 1 \le l \le r, \\ \hat{u}_k &= u_r^k. \end{aligned}$$

Again, I don't do a direct solve for $k = 1$, I just do r multigrid steps with initial guess 0. (And here I'm using the convention $I_{k-1}^k \hat{u}_{k-1} = \hat{u}_{k-1}$.)

Now we want to prove things for the $\mathcal{W}$ cycle (I don't do the $\mathcal{V}$ cycle in class).

The statement of Theorem 6.5.9 is the same, you just need a different proof for $k = 1$. Recall that $m_1 = m$, $m_2 = 0$, and $p = 2$.

But that's not hard—if you let the eigenvalues and associated eigenvectors of A_1 be λ_i, ψ_i and $\lambda_{\min} = \min_i \lambda_i$ and $\lambda_{\max} = \max_i \lambda_i$, and expand $e_0 = \sum_i \epsilon_i \psi_i$, then

$$z - MG(1, z_0, g) = \sum_i \Big(1 - \frac{\lambda_i}{\Lambda_k}\Big)^m \epsilon_i \psi_i$$

and

$$\|z - MG(1, z_0, g)\|_E^2 = \sum_i \Big(1 - \frac{\lambda_i}{\Lambda_k}\Big)^{2m} \lambda_i \epsilon_i^2 \le \Big(1 - \frac{\lambda_{\min}}{\Lambda_k}\Big)^{2m} \sum_i \lambda_i \epsilon_i^2 = \Big(1 - \frac{\lambda_{\min}}{\Lambda_k}\Big)^{2m} \|e_0\|_E^2.$$

So one need only choose m large enough that

$$\Big(1 - \frac{\lambda_{\min}}{\Lambda_k}\Big)^m \le \gamma.$$

And $\lambda_{\min}/\Lambda_k \ge C/\kappa(A_1)$, where $\kappa(A_1)$ is the condition number of A_1.

I had been wondering why the condition numbers of A_k didn't show up anywhere in the chapter, but this is why—you finesse the issue by using a direct solver in $MG(1, z_0, g)$ instead of iterating the smoother. Advantage one for your book's approach, but Theorem 6.5.9 is still true with my "lazy" approach.

At this point I used to tell the class, OK, so the proofs go through (again lazy, but now at another level!), but I realized last night that Theorem 6.7.1 does not work with my code. Let's try to go through the proof.

The problem is here:

$$\|\hat{e}_k\|_E \le \tilde{C} \Bigg\{ \gamma^r h_k |u|_2 + \gamma^{2r} h_{k-1} |u|_2 + \cdots + \gamma^{kr} h_1 |u|_2 + \left\{ \begin{array}{ll} 0, & \text{book,} \\ \gamma^{kr} \|u_1 - \hat{u}_0\|_E, & \text{code} \end{array} \right\}.$$

The problem, if you don't use a direct solver to compute $\hat{u}_1$, is that you have the extra term

$$\gamma^{kr} \|u_1\|_E = \gamma^{kr} \|P_1 u\|_E \le \gamma^{kr} \|u\|_E \le C \gamma^{kr} \|u\|_1$$

as the last term in the sum. So if $\gamma^r < 1/2$, you end up with a bound of $Ch_k(|u|_2 + \|u\|_1) = Ch_k \|u\|_2$ instead of $Ch_k |u|_2$ as the bound.

And, of course, you see this in real life. The code deals with Neumann and Robin boundary conditions, and when the solution is $u(x, y) = x + y$, for example, your multigrid solution is exact for any k and any value of the parameters (because $|u|_2 = 0$), while mine decays at an rate that is optimal.

So using a direct solver for $k = 1$ is essential for full accuracy of what you call "Full Multigrid", but not necessary for MG (but then the number of iterates m depends on $\kappa(A_1)$ in addition to whatever dependencies are in your chapter).

Anyway, it was fun to work out the consequences of my laziness.

With best wishes, Brad