James W. Lottes

Independent quality measures for optimized AMG

Mathematics and Computer Science Division
Argonne National Laboratory
9700 S Cass Avenue
Argonne
IL 60439
jlottes@mcs.anl.gov

We consider the development of an algebraic multigrid (AMG) strategy for parallel solution of linear systems with target processor counts of $P > 10^5$. We are interested in the case of many right-hand sides (e.g., as arise in the solution of time-dependent PDEs) and are therefore willing to accept significant set-up costs in favor of reduced solution time per solve. Presently, we are focusing on the case where the system of interest is the distributed coarse-grid problem that arises in domain decomposition methods or in multigrid methods that have relaxed to the coarsest-level skeleton that covers all processors. Our approach is more general, however, and we believe it would readily extend to any sparse distributed symmetric-positive definite system, coarse or fine.

Our selection of an AMG procedure is guided by the two-grid asymptotic convergence rate, equal to the spectral radius of the error propagation matrix

$$E_{\text{tg}} \equiv (I - BA)(I - PA_c^{-1}P^T A),$$

where the coarse operator is defined by $A_c \equiv P^T AP$. The iteration is determined by B, defining the *smoother*, and by the $n \times n_c$ *prolongation* matrix P. In classical AMG, the prolongation operator is further constrained in that the coarse variables are a subset of the original variables. If we order these last, then A and P take the block forms

$$A \equiv \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix}, \qquad P \equiv \begin{bmatrix} W \\ I \end{bmatrix}.$$

Thus, we may consider three components to make up a classical AMG procedure: selection of the coarse variables (*coarsening*), selection of W, and selection of B. We have come up with *independent* measures of the quality of each component, guiding our selection of each, which together imply an efficient method.

1. *Coarsening*. We propose that the coarsening ratio n_c/n and the condition number $\kappa(D_{ff}^{-1/2} A_{ff} D_{ff}^{-1/2})$ together measure the quality of a given coarsening. Here D_{ff} is the diagonal part of A_{ff}. Our simple coarsening scheme is based on Gershgorin discs and provides a guaranteed bound on κ.

2. *Prolongation weights.* We propose that the quality of W can be measured as an energy norm of the departure of W from the minimal energy weights $-A_{ff}^{-1}A_{fc}$,

$$\gamma \equiv \sup_{\mathbf{v}\neq\mathbf{0}} \frac{\|F\mathbf{v}\|_{A_{ff}}}{\|\mathbf{v}\|_{A_c}}, \qquad F \equiv W - (-A_{ff}^{-1}A_{fc}).$$

   We augment the energy-minimizing interpolation of Wan, Chan, and Smith with a dynamic procedure for determining the support of W based on a computable estimate of γ.

3. *Smoothing.* In our forthcoming paper we demonstrate that the asymptotic convergence rate is determined only by the component of B affecting the F-variables in the hierarchical basis,

$$\hat{B}_{ff} \equiv \begin{bmatrix} I & -W \end{bmatrix} B \begin{bmatrix} I & -W \end{bmatrix}^T.$$

   We propose that the quality of the smoother can be measured by the asymptotic convergence rate of the iteration using $\hat{B}_{ff}$ to solve equations governed by A_{ff},

$$\rho_f \equiv \rho(I - \hat{B}_{ff}A_{ff}).$$

   We target this measure with a simple diagonal SAI preconditioner within a Chebyshev polynomial.

The above measures are motivated by the bound, closely related to others in the literature, on the two-grid asymptotic convergence rate,

$$\rho(E_{\text{tg}}) \leq \gamma^2 + (1-\gamma^2)\rho_f,$$

which we prove in our forthcoming paper, and which immediately explains the latter two measures. The justification of the coarsening measure κ is twofold. First, a small κ implies that a few steps of diagonal preconditioning suffices to make ρ_f small. Second, a small κ implies that W may be highly sparse and yet be a good approximation to $A_{ff}^{-1}A_{fc}$, thereby inducing a small γ. This is due to a result of Demko, that the entries in the inverse of a sparse matrix (here A_{ff}^{-1}) decay exponentially with a rate characterized by the condition number. Vassilevski was the first to mention the relevance of this result to multigrid.

Although our solver is still in the development stage, we have had some initial success in deploying it on the 4096-processor BG/P at Argonne for a coarse-grid problem of dimension $n = 412,000$ (originating fine-grid problem of dimension $n_f = 44\text{M}$). The new solver yields the same outer (fine-grid) iteration count as that realized with our direct projection-based solver. After some tuning of the communication kernels the AMG solver is 6.5 times faster than the original coarse-grid solver, and the overall solution time is reduced nearly two-fold. We expect further improvements with additional tuning.