A General Strength of Connection Measure
Term Project – CS 550

Jacob Schroder*

April 21, 2007

1 Introduction

This project explores a new strength of connection measure from [2]. Strength of connection measures are crucial for the effective performance of algebraic multigrid (AMG). Strength of connection information is used to obtain good C-F splittings, to perform intelligent aggregation in the presence of anisotropies and to place nonzeros in the restriction and interpolation operators.

Ideally, AMG would be a black box solver that only requires a matrix and possibly a smoothing method. This in turn implies that the strength of connection measure should not require user defined tuning and be applicable to general matrices. Traditional strength of connection measures work well for classic sample problems such as isotropic diffusion and grid aligned anisotropic diffusion, i.e. problems that yield M-matrices. However, traditional strength of connection measures can easily fail when applied to matrices that deviate strongly from M-matrices. The strength of connection measure from [2] brings us one step closer to realizing an efficient and effective black box strength measure.

2 Preliminaries

The system under consideration is $Ax = b$.

Smooth error means algebraically smooth error, the error which is not reduced by smoothing. We can loosely define this error by first defining the smoother. For example, let

$$S = (I - \alpha D^{-1} A). \tag{1}$$

Then, we can define the smooth error as

$$Se \approx e, \tag{2}$$

---

*This work was in part supported by Sandia Corporation and Dr. Raymond Tuminaro.

with $\alpha = \frac{K}{\rho(D^{-1}A)}$ and $\rho(S) \approx 1$. The high energy eigenvectors of S define the smooth error and this space corresponds to the low energy eigenvectors of A.

A strong connection between degrees of freedom i and j means that we can reliably interpolate smooth error between them. This highlights the circular logic between interpolation and strength of connection. Defining strong connections assumes some interpolation method, and the construction of interpolation operators assumes a definition of smooth error. We will concern ourselves only with strength of connection. However, we can say that we want our interpolation basis functions to have low energy, i.e. have a small energy norm. This is a motivating idea behind smoothed aggregation. This also hints that we may want to define strength using the energy norm.

When we look at a strength of connection measure applied to an anisotropic problem, we are primarily interested in the separation inside a typical stencil of the strength of connection matrix. Separation refers to the difference in magnitude between the entries for strong and weak connections. We want a large separation as that implies a generic drop tolerance can be applied to the strength of connection matrix so that strong and weak connections are distinguished.

3 Energy Minimization Strength Measure

The eigenspace of S defines smooth error, but eigen-calculations are expensive, and it is not clear for an arbitrary operator where to truncate the eigenspace when capturing smooth error. A more accessible heuristic for strength is an approximation to A^{-1}. A^{-1} relates the size of the residual at degree of freedom, j, to the size of the error at degree of freedom, i.

$$Ae = r \Leftrightarrow A^{-1}r = e. \tag{3}$$

This is the equation that multigrid solves, so it is reasonable to think that an approximation to A^{-1} could yield useful strength information. Just such a strength measure is developed in [2].

For a strength of connection matrix, $\widetilde{A}$, the strength of connection between degrees of freedom i and j is measured by

$$\widetilde{A}_{ij} = \frac{\| \, G^{(i)} - (G^{(i)}_j \; I^{(j)}) \, \|_A}{\| \, G^{(i)} \, \|_A}, \tag{4}$$

where $I^{(j)}$ is the j-th column of the identity and $G^{(i)}$ is the i-th column of A^{-1} as approximated by some local relaxation scheme such as weighted Jacobi.

We give two motivations for this measure. The removal of an entry from $G^{(i)}$ will create a large energy norm only if its removal created large derivatives. Alternatively, imagine that degree of freedom, i, is the only coarse grid point. Then, the ideal interpolation operator would be a column of the inverse. This choice gives us a (near) minimal energy norm of our only coarse grid basis function. If the removal of point j from $(A^{-1})_{:,i}$ creates large values in the energy-norm, then points i and j are probably strongly connected, because point j's removal greatly disrupted the low energy state of $(A^{-1})_{:,i}$.

The implementation algorithm of this measure is as follows for a matrix, A, the relaxation count μ, and g, the current column of G.

```
$\omega = \frac{1}{\rho(A)}$
for i= 1 :  numRows(A)
    g = 0
    for j= 1 :  $\mu$
        g = g + $\omega$*($D^{-1}$(I(:,i) - A*g))
    end
    jj = find(A(i,:));
    denominator = sqrt(g'*A*g);
    for j = 1:length(jj)
        if(jj(j) != i)
            temp = g(jj(j));
            g(jj(j)) = 0;
            $\widetilde{A}$(i,jj(j)) = ( sqrt(g'*A*g)/denominator ) - 1;
            g(jj(j)) = temp;
        end
    end
end
Apply drop tolerance on $\widetilde{A}$ row-wise
```

We only calculate $\widetilde{A}$ for the nonzero pattern of A excluding the diagonal. We did this to limit the number of calculations and because strength information about one's algebraic neighbors should be enough to aggregate. The two user defined parameters are, μ, the number of smoothing steps and the drop tolerance. The application of the drop tolerance compares values in a row of $\widetilde{A}$ and drops those values less than a fraction of the largest value in that row. Those values not dropped are considered strong connections.

4 Empirical Results

Some empirical results were generated by using $\widetilde{A}$ inside of a smoothed aggregation based multigrid code provided by Dr. Ray Tuminaro. All operators used in computational experiments were for bilinear Q1 elements on a regular grid of the unit box. All operators were also non M-matrices. Two grid multigrid cycles were used to generate preconditioners for PCG.

4.1 Coefficient Jump

Tests were run on operators for the PDE

$$\nabla \cdot (a(x,y)\nabla u) = f, \tag{5}$$

with $a(x,y) = 1$ if $x \geq 0.5$ and $a = 100$ if $x < 0.5$. No aggregates crossed the interface boundary. Convergence counts for PCG were steady at about 8 for grid sizes of 17×17, 33×33, 65×65 and 129×129. Operator complexity stayed constant at about 1.12.

Stencils were recorded in table (1) in order to examine the separation in $\widetilde{A}$. Stencils for grid sizes 17×17, 65×65 and 129×129 were also recorded, but are not shown here as they are comparable.

μ	Point	Stencil		
4	(17,8)	0.0356	0.0240	0.0004
		0.0469	0.0000	0.0005
		0.0356	0.0240	0.0004
4	(18,8)	0.0003	0.0149	0.0160
		0.0005	0.0000	0.0211
		0.0003	0.0149	0.0160

Table 1: $\widetilde{A}$ stencils around interface – Grid Size is 33×33.

Strong connections are correctly discerned by the measure because only the connections between points where the PDE has the same the coefficient are strong. Point (17,8) is the first point on the left side of the interface and is correctly connected strongly to its neighbors on the left, above and below. Point (18,8) is the first point on the right side of the interface and is correctly connected strongly to its neighbors on the right, above and below.

4.2 Rotated Anisotropy

Tests were also run on operators for rotated anisotropy,

$$-(\nu c^2 + s^2)u_{xx} + 2(1-\nu)sc\, u_{xy} - (\nu s^2 + c^2)u_{yy} = f. \tag{6}$$

$c = cos(\frac{\pi}{4})$, $s = sin(\frac{\pi}{4})$ and $\nu = 0.001$. Aggregates lined up nicely in a 45° angle. Convergence counts for PCG were steady at about 10 for grid sizes 17×17, 33×33, 65×65 and 129×129. Operator complexity was constant at about 1.52.

Stencils were recorded in table (2) in order to examine the separation in $\widetilde{A}$. Stencils for grid sizes of 17×17, 33×33, 65×65 and 129×129 were also recorded, but are comparable. The flow of information at a 45° angle is captured well by this measure. The separation is about 1 order of magnitude greater in the direction of strongest flow, from bottom left to upper right, when compared to the direction of weakest flow, from upper left to bottom right.

μ	Stencil		
2	0.0011	0.0073	0.0378
	0.0073	0.0000	0.0073
	0.0378	0.0073	0.0011
3	0.0012	0.0151	0.0634
	0.0151	0.0000	0.0151
	0.0634	0.0151	0.0012
4	0.0010	0.0230	0.0840
	0.0230	0.0000	0.0230
	0.0840	0.0230	0.0010

Table 2: $\widetilde{A}$ stencil at point (17,8) – Grid Size is 33×33.

4.3 Cost

One of the main drawbacks to this measure is the cost. Consider the calculation of one column, $G^{(i)}$, of the approximate inverse, G.

$$\begin{aligned} G_0^{(i)} &= 0 \qquad (7) \\ G_1^{(i)} &= 0 + \omega D^{-1}(e_j - 0) = \omega D^{-1} e_j \\ G_2^{(i)} &= \omega D^{-1} e_j + \omega D^{-1}(e_j - A\omega D^{-1} e_j), \\ &\quad \text{which is } O(D^{-1}AD^{-1}e_j). \end{aligned}$$

Considering just this dominant operation, the following pattern becomes clear.

$$\begin{aligned} G_3^{(i)} &= O(D^{-1}AD^{-1}AD^{-1}e_j) = O(A^2 e_j) \qquad (8) \\ G_4^{(i)} &= O(D^{-1}A(D^{-1}AD^{-1}AD^{-1}e_j) = O(A^3 e_j) \end{aligned}$$

If one considers this operation for each column of the identity, then the cost of doing k Jacobi iterations is of the same order as k-1 matrix multiplies with A. However, there is an additional A introduced when the measure, (4), is computed. Both the numerator and denominator require an A-norm of $G^{(i)}$. Hence, the cost of computing the measure, (4), when using k smoothing steps is of the same order as k matrix multiplies with A.

4.4 Using A^{-1}

We also calculated $\widetilde{A}$ using actual columns of the inverse. The results were not encouraging. Not only were convergence iterations in PCG slightly worse, but as indicated in table, (3), separation inside $\widetilde{A}$ was much worse. This indicates something goes wrong in the limit as $G^{(i)}$ approaches the actual column of A^{-1}.

Stencil		
0.1282	0.4390	0.6246
0.3854	0.0000	0.3800
0.4812	0.3162	0.0712

Table 3: $\widetilde{A}$ stencil at point (17,8) – Rotated Anisotropy of $\frac{\pi}{4}$

4.5 SPAI

Perhaps, using A^{-1} with its unrestricted stencil to calculate $\widetilde{A}$ smears information, so that we lose separation. To examine this possible issue, we used SPAI [4] to calculate G. This should eliminate the smearing issue because SPAI will find a good approximate inverse restricted to a small number of nonzeros. We enforced $\text{nnz}(G) \approx \text{nnz}(A)$. Examining tables (4 - 6), we see that SPAI offers limited improvement over A^{-1} and this in turn indicates that there may be a deeper problem with the new measure. We are led empirically to the conclusion that A^{-1} does not provide ideal strength information.

Stencil		
0.0051	0.1450	0.0042
0.0095	0.0000	0.0079
0.0057	0.1599	0.0039

Table 4: $\widetilde{A}$ stencil at point (17,8) – Vertically strong anisotropy

Stencil		
0.0000	0.0374	0.0921
0.0378	0.0000	0.0285
0.0944	0.0288	0.0000

Table 5: $\widetilde{A}$ stencil at point (17,8) – $\frac{\pi}{4}$ rotated anisotropy

Stencil		
0.0040	0.0997	0.0227
0.0005	0.0000	0.0005
0.0227	0.0997	0.0040

Table 6: $\widetilde{A}$ stencil at point (17,8) – $\frac{\pi}{8}$ rotated anisotropy

5 Conclusion

5.1 Summary

The major drawbacks to the new strength measure are as follows.

1. The cost can be prohibitive. Using k Jacobi relaxations to calculate the measure is equivalent to k matrix-matrix multiplies with A.

2. The measure did not perform well for the most difficult test case, a stretched and unstructured grid on an oval-shaped domain. Yet, it did do as well as the distance-based strength of connection measure.

3. As $G \rightarrow A^{-1}$, we get poor strength information. This is because A^{-1} relates degrees of freedom with respect to all error for the system, $Ae = r$. This includes smooth error and also error effectively reduced by smoothing.

   This issue is partially avoided by the fact that we only smooth a handful of times when calculating G. After only a few iterations of smoothing, G should still be heavily influenced by the smoothing method and hence; hopefully, local and reflective of the smoothing method.

The major benefits of this measure are as follows.

1. The proposed measure in [2] is effective for a more general class of matrices than the classic strength of connection measure.

2. It avoids any sort of randomness in its calculation, unlike CR.

3. It is not an algorithm that requires much tuning. Small values of $\mu = 2, 3$ and 4 worked well with a drop tolerance of 0.25.

4. It works well in practice for small μ on a variety of test problems. It is encouraging that it works as well as the distance based strength of connection measure for grid induced anisotropies. This implies that it is a good strength measure for non-grid induced anisotropies that are nonetheless equivalent to grid induced anisotropies

5.2 Future Ideas

1. Reduce the cost of the measure by some further approximation or by only calculating the measure for rows that deviate strongly from being M-like.

2. Explore adaptive threshholding by varying μ.

3. Explore other strength measures that are based on approximating the smooth components of S's eigenspace. This approach is an area of my active work and attacks the problem of strong connections at its roots.

6 References

1. A. Brandt, S. McCormick and J. W. Ruge. *Algebraic multigrid (AMG) for sparse matrix equations.* Sparsity and Its Applications, DJ Evans, ed., Cambridge University Press, Cambridge, 1984.

2. J. Brannick, M. Brezina, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge. *An energy-based AMG coarsening strategy.* Num. Lin. Alg. Appl., 2006: **13**:133-148.

3. J. W. Ruge and K. Stüben. *Algebraic Multigrid,* in Multigrid Methods. S. McCormick, ed., SIAM, Philadelphia, 1987, pp. 73-130.

4. M. Grote and T. Huckle. *Parallel preconditioning with sparse approximate inverses.* SIAM J. Sci. Comput., 1997: **18**:838-853.