

Bootstrap AMG for Markov Chain Computations

Karsten Kahl

Bergische Universität Wuppertal

May 27, 2010

Outline

Markov Chains

Algebraic Multigrid

Subspace Eigenvalue Approximation

Ingredients of Algebraic Multigrid

Least Squares Interpolation

Algebraic Multigrid Eigensolver — Bootstrap Setup

Numerical Results

Conclusion and Outlook



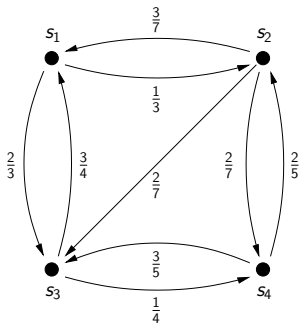
Finite homogeneous Markov Chains

- ▶ States $s_1, \dots, s_n, n \in \mathbb{N}$
- ▶ Transition probabilities $p(s_j \rightarrow s_i)$
- ⇒ Transition matrix $T = (t_{ij})_{i,j} \in \mathbb{R}^{n \times n}$

$$t_{ij} = p(s_j \rightarrow s_i)$$

- ▶ Example:

$$T = \begin{pmatrix} 0 & \frac{3}{7} & \frac{3}{4} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{2}{5} \\ \frac{2}{7} & \frac{2}{7} & 0 & \frac{3}{5} \\ 0 & \frac{2}{7} & \frac{1}{4} & 0 \end{pmatrix}$$





Properties of Markov Chains

- ▶ Column-stochastic transition matrices T ,

$$\mathbf{1}^t T = \mathbf{1}^t, \quad \mathbf{1}^t = (1 \quad 1 \quad \dots \quad 1)$$

- ▶ T irreducible \Rightarrow existence of unique steady-state solution

$$Tx = x, \quad x > 0, \quad \|x\|_1 = 1$$

- ▶ $\text{spec}_r(B) \subset \mathcal{U}_1(0) = \{z \in \mathbb{C}, |z| < 1\}$
- ▶ Equivalent formulation as system of linear equations

$$Ax = (I - T)x = 0$$

- ▶ $\mathbf{1}^t A = 0$, $\text{span}(\{x\}) = \text{null}_r(A)$, $\text{spec}_r(A) \subset \mathcal{U}_1(1)$



Power Method

- ▶ Iterative method for computation of steady-state solution

$$x^{k+1} = Tx^k = (I - A)x^k = x^k - Ax^k$$

- ▶ Power method for $T \Leftrightarrow$ Richardson iteration for A
- ▶ With $1 = \lambda_1 > |\lambda_2| \geq \dots \geq |\lambda_n| \geq 0$ eigenvalues of T

$$\|x^k - x\|_2 = c|\lambda_2|^k$$

\Rightarrow Slow convergence for $\lambda_2 \approx 1$, yet fast reduction of components of x^k corresponding to $\lambda_i \ll 1$

Approximation of Eigenpair $Tx = \lambda_{max}x$ in a subspace \mathcal{V}

Basis v_1, \dots, v_k of \mathcal{V} , $V = (v_1 \ \dots \ v_k)$

- ▶ Galerkin formulation

$$V^t T V y = \tilde{\lambda}_{max} V^t V y \quad \Rightarrow \quad \tilde{x} = V y \approx x, \quad \tilde{\lambda}_{max} \approx \lambda_{max}$$

- ▶ Petrov-Galerkin formulation

$$W^t T V y = \tilde{\lambda}_{max} W^t V y \quad \Rightarrow \quad \tilde{x} = V y \approx x, \quad \tilde{\lambda}_{max} \approx \lambda_{max}$$

Special case: Krylov subspaces $\mathcal{V} = \mathcal{K}_k$

- ▶ Orthonormal basis v_1, \dots, v_k , $V^t V = I$
- ▶ Bi-orthogonal basis v_1, \dots, v_k and w_1, \dots, w_k , $W^t V = I$



Ingredients of Algebraic Multigrid (Needed, Given)

- ▶ Sparse linear system of equations $Au = f$
- ▶ Hierarchy of sparse systems of linear equations $A_l u_l = f_l$
- ▶ Appropriate smoothing iterations S_l
- ▶ Definition of restriction and interpolation

$$R_l^{l+1} : V_l \rightarrow V_{l+1} \quad \text{and} \quad P_{l+1}^l : V_{l+1} \rightarrow V_l$$

- ▶ Definition of operator hierarchy $A = A_0, A_1, \dots, A_L$

$$A_{l+1} = R_l^{l+1} A_l P_{l+1}^l$$

⇒ Columns of P_{l+1}^l form sparse basis for subspace V_{l+1} in V_l complementary to the smoother



Least Squares Interpolation

Computation of interpolation weights $(p_i)_j$, $i \in \mathcal{F}, j \in \mathcal{C}_i$

$$\mathcal{L}_{\mathcal{C}_i}(p_i) = \sum_{s=1}^k \omega_s (u_i^{(s)} - \sum_{j \in \mathcal{C}_i} (p_i)_j u_j^{(s)})^2 \rightarrow \min_{p_i}$$

- ▶ Test vectors $u^{(1)}, \dots, u^{(k)} \in \mathbb{R}^n$,
- ▶ Set of interpolatory points $\mathcal{C}_i \subset \mathcal{C}$ for $i \in \mathcal{F}$
- ▶ Weights $\omega_s \in \mathbb{R}^+$ to bias LS fit, e.g., $\omega_s = \left(\frac{\|u^{(s)}\|}{\|Au^{(s)}\|} \right)^2$
- ▶ Splitting of variables $\Omega = \mathcal{F} \cup \mathcal{C}$, Interpolation P from \mathcal{C} to Ω ,

$$P = \begin{pmatrix} P_{fc} \\ I \end{pmatrix}, \quad p_{ij} \neq 0, i \in \mathcal{F}, j \in \mathcal{C}_i$$



Algebraic Multigrid Eigensolver — Bootstrap Setup

- ▶ Test vectors initially positive random, slightly smoothed

$$S^\eta u^{(1)}, \dots, S^\eta u^{(k)}$$

- ▶ MG hierarchy with interpolation P_{l+1}^l , averaging restriction R_l^{l+1} ($\mathbf{1}^t R_l^{l+1} = \mathbf{1}^t$)

$$A_{l+1} = R_l^{l+1} A_l P_{l+1}^l$$

$$B_{l+1} = R_l^{l+1} B_l P_{l+1}^l$$

- ▶ New approximation of steady-state vector, enhancement of TV set, quality control of MG hierarchy

Bootstrap Multigrid Eigensolver

if Coarsest level **then**

$$\mathbf{V}_L = \{v_i^{(L)}\}_{i=1, \dots, k_v}$$

$$A_L v_i^{(L)} = \lambda_i^{(L)} B_L v_i^{(L)}$$

else

$$v_i^{(l)} = P_{l+1}^l v_i^{(l+1)}$$

$$\lambda_i^{(l)} = \lambda_i^{(l+1)}$$

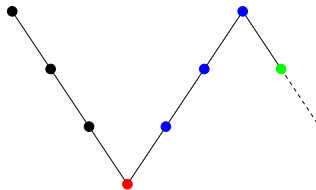
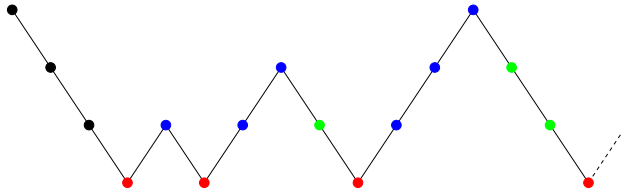
$$\text{Relax } \left(A_l - \lambda_i^{(l)} B_l \right) v_i^{(l)} = 0$$

$$\text{Update } \lambda_i^{(l)} = \frac{\langle A_l v_i^{(l)}, v_i^{(l)} \rangle_2}{\langle B_l v_i^{(l)}, v_i^{(l)} \rangle_2}$$

end if

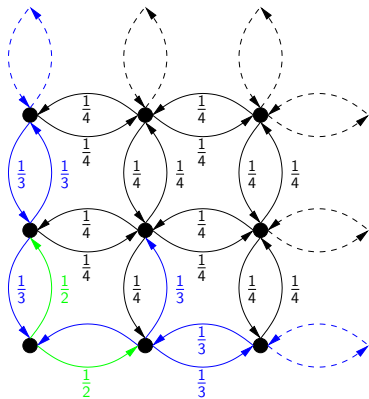


Bootstrap Multigrid Eigensolver Cycling Strategies



- Relax on $Au = 0, u \in \mathbf{U}$
- Compute \mathbf{V} , s.t., $Av = \lambda Bv, v \in \mathbf{V}$
- Relax on $Av = \lambda Bv, v \in \mathbf{V}$
- Relax on $Au = 0, u \in \mathbf{U}$ and $Av = \lambda Bv, v \in \mathbf{V}$

Uniform 2D Network

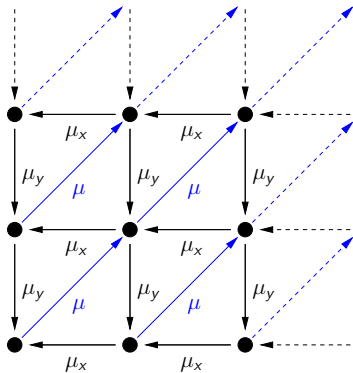


N	17	33	65	129
MLE	10	9	10	11
pArnoldi	7_V	8_V	10_V	9_{V^2}

- ▶ Accuracy 10^{-8}
- ▶ ω -Jacobi smoother with $\omega = .7$
- ▶ TV sets $|\mathbf{U}| = |\mathbf{V}| = 6$
- ▶ V(2,2)-cycle, caliber 2
- ▶ Coarsest grid 5×5



Tandem Queueing Network

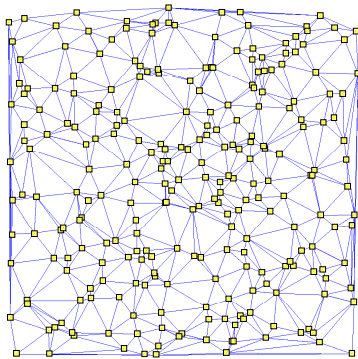


$$\mu = \frac{11}{31}, \quad \mu_x = \frac{10}{31}, \quad \mu_y = \frac{10}{31}$$

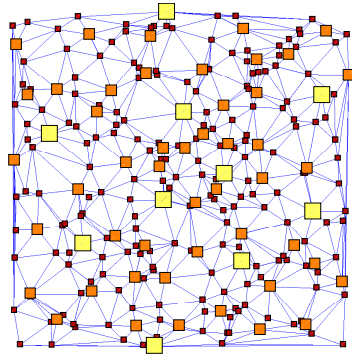
N	17	33	65	129
MLE	8	8	8	8
pArnoldi	6_V	6_V	6_V	7_V

- ▶ Accuracy 10^{-8}
- ▶ ω -Jacobi smoother with $\omega = .7$
- ▶ TV sets $|\mathbf{U}| = |\mathbf{V}| = 6$
- ▶ V(2,2)-cycle, caliber 2
- ▶ Coarsest grid 5×5

Random Planar Graph



(a) Random planar graph, $n = 256$



(b) Coarsening, $n = 256$

$$p(j \rightarrow i) = d_{out}(j)^{-1}$$



Random Planar Graph

N	256	512	1024	2048
MLE	15	20	20	20
pArnoldi	8_V	10_V	10_V	11_V

- ▶ Accuracy 10^{-8}
- ▶ ω -Jacobi smoother with $\omega = .7$
- ▶ TV sets $|\mathbf{U}| = |\mathbf{V}| = 6$, caliber 2
- ▶ $V(2, 2)$ -cycle, 3-grid method



Conclusions and Outlook

Conclusions

- ▶ Algebraic Multigrid Eigensolver
- ▶ Reuse of AMG hierarchy as a “preconditioner”
- ▶ Promising and scaling results for simple Markov Chains

Outlook

- ▶ More complex Markov Chains
- ▶ Theory for LS interpolation and MG Eigensolver
- ▶ Efficient (parallel) implementation of Bootstrap AMG



Thank you for your attention!