Updating Markov Chains

Carl Meyer
Amy Langville

Department of Mathematics
North Carolina State University
Raleigh, NC

A. A. Markov Anniversary Meeting

June 13, 2006

Intro

Assumptions

Very large irreducible chain

- $m = O(\mathbf{10^9})$ number of states
- $\mathbf{Q}_{m\times m}$ old transition matrix
- $\boldsymbol{\phi}^T = (\phi_{\mathbf{1}}, \phi_{\mathbf{2}}, \ldots, \phi_m)$ old stationary distribution

Intro

Assumptions

Very large irreducible chain

- $m = O(10^9)$ number of states
- $\mathbf{Q}_{m \times m}$ old transition matrix
- $\phi^T = (\phi_1, \phi_2, \ldots, \phi_m)$ old stationary distribution

Updates

Change some transition probabilities

Add or delete some states

- $\mathbf{P}_{n \times n}$ new transition matrix (irreducible)
- $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \ldots, \pi_n)$ new distribution (unknown)

Intro

Assumptions

Very large irreducible chain

- $m = O(10^9)$ number of states
- $\mathbf{Q}_{m\times m}$ old transition matrix
- $\boldsymbol{\phi}^T = (\phi_1, \phi_2, \ldots, \phi_m)$ old stationary distribution

Updates

Change some transition probabilities

Add or delete some states

- $\mathbf{P}_{n\times n}$ new transition matrix (irreducible)
- $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \ldots, \pi_n)$ new distribution (unknown)

Aim

Use $\boldsymbol{\phi}^T$ to compute $\boldsymbol{\pi}^T$

Exact (Theoretical) Updating

Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \quad \Longrightarrow \quad \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T$$

Exact (Theoretical) Updating

Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \implies \pi^T = \phi^T - \epsilon^T \qquad \epsilon^T = \phi^T \mathbf{EZ}(\mathbf{I} + \mathbf{EZ})^{-1}$$

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^{\#} \quad \text{(group inverse)} \end{cases}$$

Exact (Theoretical) Updating

Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \quad \Longrightarrow \quad \boldsymbol{\pi}^T = \phi^T - \epsilon^T \qquad \epsilon^T = \phi^T \mathbf{EZ}(\mathbf{I} + \mathbf{EZ})^{-1}$$

Requires $m = n$ (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^{\#} \quad \text{(group inverse)} \end{cases}$$

Exact (Theoretical) Updating

Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \quad \Longrightarrow \quad \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T \qquad \boldsymbol{\epsilon}^T = \boldsymbol{\phi}^T \mathbf{EZ}(\mathbf{I} + \mathbf{EZ})^{-1}$$

Requires $m = n$ (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^{\#} \quad \text{(group inverse)} \end{cases}$$

One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \boldsymbol{\delta}_i^T \quad \Longrightarrow \quad \boldsymbol{\pi}^T = \boldsymbol{\phi}^T - \boldsymbol{\epsilon}^T$$

Exact (Theoretical) Updating

Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \quad \Longrightarrow \quad \pi^T = \phi^T - \epsilon^T \qquad \epsilon^T = \phi^T \mathbf{EZ}(\mathbf{I} + \mathbf{EZ})^{-1}$$

Requires $m = n$ (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^{\#} \quad \text{(group inverse)} \end{cases}$$

One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \boldsymbol{\delta}_i^T \quad \Longrightarrow \quad \boldsymbol{\pi}^T = \phi^T - \epsilon^T$$

$$\epsilon^T = \left(\frac{\phi_i}{1 + \boldsymbol{\delta}_i^T \mathbf{A}_{*i}^{\#}} \right) \boldsymbol{\delta}^T \mathbf{A}^{\#} \qquad \mathbf{A}^{\#} = (\mathbf{I} - \mathbf{Q})^{\#}$$

Exact (Theoretical) Updating

Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \quad \Longrightarrow \quad \pi^T = \phi^T - \epsilon^T \qquad \epsilon^T = \phi^T \mathbf{EZ}(\mathbf{I} + \mathbf{EZ})^{-1}$$

Requires $m = n$ (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^{\#} \quad \text{(group inverse)} \end{cases}$$

One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \delta_i^T \quad \Longrightarrow \quad \pi^T = \phi^T - \epsilon^T$$

$$\epsilon^T = \left(\frac{\phi_i}{1 + \delta_i^T \mathbf{A}_{*i}^{\#}} \right) \delta^T \mathbf{A}^{\#} \qquad \mathbf{A}^{\#} = (\mathbf{I} - \mathbf{Q})^{\#}$$

$$(\mathbf{I} - \mathbf{P})^{\#} = \mathbf{A}^{\#} + \mathbf{e}\epsilon^T \left[\mathbf{A}^{\#} - \gamma \mathbf{I} \right] - \frac{\mathbf{A}_{*i}^{\#} \epsilon^T}{\phi_i} \qquad \gamma = \frac{\epsilon^T \mathbf{A}_{*i}^{\#}}{\phi_i}$$

Exact (Theoretical) Updating

Perturbation Formula

$$\mathbf{P} = \mathbf{Q} - \mathbf{E} \quad \Longrightarrow \quad \pi^T = \phi^T - \epsilon^T \qquad \epsilon^T = \phi^T \mathbf{EZ}(\mathbf{I} + \mathbf{EZ})^{-1}$$

Requires $m = n$ (No states added or deleted)

$$\mathbf{Z} = \begin{cases} \text{Fundamental Matrix} \\ (\mathbf{I} - \mathbf{Q})^\# \quad \text{(group inverse)} \end{cases}$$

One Row At A Time (Sherman–Morrison)

$$\mathbf{p}_i^T = \mathbf{q}_i^T - \delta_i^T \quad \Longrightarrow \quad \pi^T = \phi^T - \epsilon^T$$

$$\epsilon^T = \left(\frac{\phi_i}{1 + \delta_i^T \mathbf{A}^\#_{*i}} \right) \delta^T \mathbf{A}^\# \qquad \mathbf{A}^\# = (\mathbf{I} - \mathbf{Q})^\#$$

$$(\mathbf{I} - \mathbf{P})^\# = \mathbf{A}^\# + \mathbf{e}\epsilon^T \left[\mathbf{A}^\# - \gamma \mathbf{I} \right] - \frac{\mathbf{A}^\#_{*i} \epsilon^T}{\phi_i} \qquad \gamma = \frac{\epsilon^T \mathbf{A}^\#_{*i}}{\phi_i}$$

Not Practical For Large Problems

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

— Requires $m = n$ (no states added or deleted)

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

— Requires $m = n$ (no states added or deleted)

— Requires $\phi^T \approx \boldsymbol{\pi}^T$ (generally means $\mathbf{P} \approx \mathbf{Q}$)

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

- Requires $m = n$ (no states added or deleted)
- Requires $\phi^T \approx \pi^T$ (generally means $\mathbf{P} \approx \mathbf{Q}$)
- Asymptotic rate of convergence: $R = -\log_{10} |\lambda_2|$

  Need about $1/R$ iterations to eventually gain one additional significant digit of accuracy

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

— Requires $m = n$ (no states added or deleted)

— Requires $\phi^T \approx \pi^T$ (generally means $\mathbf{P} \approx \mathbf{Q}$)

— Asymptotic rate of convergence: $R = -\log_{10} |\lambda_2|$

Need about $1/R$ iterations to eventually gain one additional significant digit of accuracy

Example:

- ▷ Suppose $digit_1(\phi^T) = digit_1(\boldsymbol{\pi}^T)$
- ▷ Want 12 digits of accuracy

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

— Requires $m = n$ (no states added or deleted)

— Requires $\phi^T \approx \pi^T$ (generally means $\mathbf{P} \approx \mathbf{Q}$)

— Asymptotic rate of convergence: $R = -\log_{10} |\lambda_2|$

Need about $1/R$ iterations to eventually gain one additional significant digit of accuracy

Example:

- ▷ Suppose $digit_1(\phi^T) = digit_1(\pi^T)$
- ▷ Want 12 digits of accuracy
- ▷ RPM — about $11/R$ iterations
- ▷ Start from scratch — about $12/R$ iterations

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

- Requires $m = n$ (no states added or deleted)
- Requires $\phi^T \approx \pi^T$ (generally means $\mathbf{P} \approx \mathbf{Q}$)
- Asymptotic rate of convergence: $R = -\log_{10} |\lambda_2|$

  Need about $1/R$ iterations to eventually gain one additional significant digit of accuracy

Example:

- Suppose $digit_1(\phi^T) = digit_1(\pi^T)$
- Want 12 digits of accuracy
- RPM — about $11/R$ iterations
- Start from scratch — about $12/R$ iterations
- Only 8.3% reduction

Restarted Power Method

$\mathbf{x}_{j+1}^T = \mathbf{x}_j^T \mathbf{P}$ **with** $\mathbf{x}_0^T = \phi^T$ (assume aperiodic)

— Requires $m = n$ (no states added or deleted)

— Requires $\phi^T \approx \boldsymbol{\pi}^T$ (generally means **P** $\approx$ **Q**)

— Asymptotic rate of convergence: $R = -\log_{10} |\lambda_2|$

Need about $1/R$ iterations to eventually gain one additional significant digit of accuracy

Example:

- ▷ Suppose $digit_1(\phi^T) = digit_1(\boldsymbol{\pi}^T)$
- ▷ Want 12 digits of accuracy
- ▷ RPM — about $11/R$ iterations
- ▷ Start from scratch — about $12/R$ iterations
- ▷ Only 8.3% reduction

A Little Better, But Not Great

Censoring

Partition (not necessarily NCD)

$$
\mathbf{P}_{n\times n} = \begin{array}{c} \\ G_1 \\ G_2 \\ \vdots \\ G_k \end{array} \!\!\! \begin{array}{c} \begin{array}{cccc} G_1 & G_2 & \cdots & G_k \end{array} \\ \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1k} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{k1} & \mathbf{P}_{k2} & \cdots & \mathbf{P}_{kk} \end{pmatrix} \end{array}
$$

Censoring

Partition (not necessarily NCD)

$$\mathbf{P}_{n\times n} = \begin{array}{c} \\ G_1 \\ G_2 \\ \vdots \\ G_k \end{array} \begin{array}{c} \begin{array}{cccc} G_1 & G_2 & \cdots & G_k \end{array} \\ \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1k} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{k1} & \mathbf{P}_{k2} & \cdots & \mathbf{P}_{kk} \end{pmatrix} \end{array}$$

Censored Chains

Records state of process only when chain visits states in G_i.

Visits to states outside of G_i are ignored.

Censoring

Partition (not necessarily NCD)

$$\mathbf{P}_{n\times n} = \begin{array}{c} \\ G_1 \\ G_2 \\ \vdots \\ G_k \end{array} \begin{array}{c} \begin{array}{cccc} G_1 & G_2 & \cdots & G_k \end{array} \\ \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1k} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{k1} & \mathbf{P}_{k2} & \cdots & \mathbf{P}_{kk} \end{pmatrix} \end{array}$$

Censored Chains

Records state of process only when chain visits states in G_i.

Visits to states outside of G_i are ignored.

Censored Transition Matrices

$$\mathbf{C}_i = \mathbf{P}_{ii} + \mathbf{P}_{i\star}(\mathbf{I} - \mathbf{P}_i^\star)^{-1}\mathbf{P}_{\star i}$$

Stochastic Complements

Aggregation

Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$

Aggregation

Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$

Aggregation Matrix

$$\mathcal{A}_{k\times k} = \begin{bmatrix} \mathbf{s}_1^T \mathbf{P}_{11}\mathbf{e} & \cdots & \mathbf{s}_1^T \mathbf{P}_{1k}\mathbf{e} \\ \vdots & \ddots & \vdots \\ \mathbf{s}_k^T \mathbf{P}_{k1}\mathbf{e} & \cdots & \mathbf{s}_k^T \mathbf{P}_{kk}\mathbf{e} \end{bmatrix} \qquad \mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Aggregation

Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$

Aggregation Matrix

$$\mathcal{A}_{k\times k} = \begin{bmatrix} \mathbf{s}_1^T \mathbf{P}_{11}\mathbf{e} & \cdots & \mathbf{s}_1^T \mathbf{P}_{1k}\mathbf{e} \\ \vdots & \ddots & \vdots \\ \mathbf{s}_k^T \mathbf{P}_{k1}\mathbf{e} & \cdots & \mathbf{s}_k^T \mathbf{P}_{kk}\mathbf{e} \end{bmatrix} \qquad \mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Aggregated Distribution

$$\boldsymbol{\alpha}^T \mathcal{A} = \boldsymbol{\alpha}^T \qquad \boldsymbol{\alpha}^T = (\alpha_1, \alpha_2, \ldots, \alpha_k)$$

Aggregation

Censored Distributions

$$\mathbf{s}_i^T \mathbf{C}_i = \mathbf{s}_i^T$$

Aggregation Matrix

$$\mathcal{A}_{k\times k} = \begin{bmatrix} \mathbf{s}_1^T \mathbf{P}_{11}\mathbf{e} & \cdots & \mathbf{s}_1^T \mathbf{P}_{1k}\mathbf{e} \\ \vdots & \ddots & \vdots \\ \mathbf{s}_k^T \mathbf{P}_{k1}\mathbf{e} & \cdots & \mathbf{s}_k^T \mathbf{P}_{kk}\mathbf{e} \end{bmatrix} \qquad \mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Aggregated Distribution

$$\alpha^T \mathcal{A} = \alpha^T \qquad \boldsymbol{\alpha}^T = (\alpha_1, \alpha_2, \ldots, \alpha_k)$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = (\boldsymbol{\pi}_1^T \mid \boldsymbol{\pi}_2^T \mid \cdots \mid \boldsymbol{\pi}_k^T) = (\alpha_1 \mathbf{s}_1^T \mid \alpha_2 \mathbf{s}_2^T \mid \cdots \mid \alpha_2 \mathbf{s}_2^T)$$

Partitioning

Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.

Partitioning

Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.

State Space Partition

$$\mathcal{S} = G \cup \overline{G} \qquad\qquad G = \begin{cases} \text{states likely to be most affected} \\ \text{newly added states} \end{cases}$$

Partitioning

Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.

State Space Partition

$\mathcal{S} = G \cup \overline{G}$ $\qquad$ $G = \begin{cases} \text{states likely to be most affected} \\ \text{newly added states} \end{cases}$

- $g = |G| << |\overline{G}|$

Partitioning

Intuition

Update relatively small number of states in large sparse chain

- Effects are primarily local
- Most stationary probabilities not significantly affected.

State Space Partition

$$\mathcal{S} = G \cup \overline{G} \qquad G = \begin{cases} \text{states likely to be most affected} \\ \text{newly added states} \end{cases}$$

- $g = |G| << |\overline{G}|$

Induced Matrix Partition

$$\mathbf{P}_{n\times n} = \begin{matrix} & G & \overline{G} \\ G & \mathbf{P}_{11} & \mathbf{P}_{12} \\ \overline{G} & \mathbf{P}_{21} & \mathbf{P}_{22} \end{matrix} = \left[\begin{array}{c|c|c|c} p_{11} & \cdots & p_{1g} & \mathbf{P}_{1\star} \\ \hline \vdots & & \vdots & \vdots \\ \hline p_{g1} & \cdots & p_{gg} & \mathbf{P}_{g\star} \\ \hline \mathbf{P}_{\star 1} & \cdots & \mathbf{P}_{\star g} & \mathbf{P}_{22} \end{array}\right]$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T\mathbf{C}_{g+1}$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T\mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T\mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Aggregated Distribution

$$\boldsymbol{\alpha}^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = \left(\pi_1, \ldots, \pi_g \,|\, \overline{\boldsymbol{\pi}}^T\right) = (\alpha_1, \ldots, \alpha_g, \,|\, \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Aggregated Distribution

$$\alpha^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = \left(\pi_1, \ldots, \pi_g \,|\, \overline{\boldsymbol{\pi}}^T\right) = (\alpha_1, \ldots, \alpha_g, \,|\, \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Old Distribution (reorded)

$$\phi^T = (\phi_1, \phi_2, \ldots \,|\, \overline{\phi}^T)$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Aggregated Distribution

$$\alpha^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = \left(\pi_1, \ldots, \pi_g \,|\, \overline{\boldsymbol{\pi}}^T\right) = (\alpha_1, \ldots, \alpha_g, \,|\, \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Old Distribution (reorded)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \ldots \,|\, \overline{\boldsymbol{\phi}}^T)$$

The Assumption

$$\overline{\boldsymbol{\phi}}^T \approx \overline{\boldsymbol{\pi}}^T$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T\mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T\mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Aggregated Distribution

$$\alpha^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = \left(\pi_1, \ldots, \pi_g \,\middle|\, \overline{\boldsymbol{\pi}}^T\right) = (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Old Distribution (reorded)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \ldots \mid \overline{\boldsymbol{\phi}}^T)$$

The Assumption

$$\overline{\boldsymbol{\phi}}^T \approx \overline{\boldsymbol{\pi}}^T \quad \Longrightarrow \quad \overline{\boldsymbol{\phi}}^T \approx \alpha_{g+1}\mathbf{s}_{g+1}^T$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Aggregated Distribution

$$\alpha^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = \left(\pi_1, \ldots, \pi_g \,\middle|\, \overline{\boldsymbol{\pi}}^T\right) = (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Old Distribution (reorded)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \ldots \mid \overline{\boldsymbol{\phi}}^T)$$

The Assumption

$$\overline{\boldsymbol{\phi}}^T \approx \overline{\boldsymbol{\pi}}^T \implies \overline{\boldsymbol{\phi}}^T \approx \alpha_{g+1}\mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\overline{\boldsymbol{\phi}}^T}{\overline{\boldsymbol{\phi}}^T \mathbf{e}}$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T\mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T\mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Aggregated Distribution

$$\alpha^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\pi^T = (\pi_1, \ldots, \pi_g \mid \overline{\pi}^T) = (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Old Distribution (reorded)

$$\phi^T = (\phi_1, \phi_2, \ldots \mid \overline{\phi}^T)$$

The Assumption

$$\overline{\phi}^T \approx \overline{\pi}^T \implies \overline{\phi}^T \approx \alpha_{g+1}\mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\overline{\phi}^T}{\overline{\phi}^T\mathbf{e}}$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

Aggregation Matrix

Aggregated Distribution

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)} \longrightarrow \boldsymbol{\alpha}^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = (\pi_1, \ldots, \pi_g \mid \overline{\boldsymbol{\pi}}^T) = (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Old Distribution (reorded)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \ldots \mid \overline{\boldsymbol{\phi}}^T)$$

The Assumption

$$\overline{\boldsymbol{\phi}}^T \approx \overline{\boldsymbol{\pi}}^T \implies \overline{\boldsymbol{\phi}}^T \approx \alpha_{g+1}\mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\overline{\boldsymbol{\phi}}^T}{\overline{\boldsymbol{\phi}}^T \mathbf{e}}$$

Specialized Aggregation

Censored Transition Matrices

$$\mathbf{C}_1 = \cdots = \mathbf{C}_g = [1]_{1\times 1} \qquad \mathbf{C}_{g+1} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$$

Censored Distributions

$$\mathbf{s}_1^T = \cdots = \mathbf{s}_g^T = 1 \qquad \mathbf{s}_{g+1}^T = \mathbf{s}_{g+1}^T \mathbf{C}_{g+1}$$

Aggregation Matrix

$$\mathcal{A} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}_{g+1}^T \mathbf{P}_{21} & 1 - \mathbf{s}_{g+1}^T \mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Aggregated Distribution

$$\boldsymbol{\alpha}^T = (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Aggregation Theorem

$$\boldsymbol{\pi}^T = (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}_{g+1}^T)$$

Updated Distribution

Old Distribution (reorded)

$$\boldsymbol{\phi}^T = (\phi_1, \phi_2, \ldots \mid \overline{\boldsymbol{\phi}}^T)$$

The Assumption

$$\overline{\boldsymbol{\phi}}^T \approx \overline{\boldsymbol{\pi}}^T \implies \overline{\boldsymbol{\phi}}^T \approx \alpha_{g+1}\mathbf{s}_{g+1}^T \implies \mathbf{s}_{g+1}^T \approx \frac{\overline{\boldsymbol{\phi}}^T}{\overline{\boldsymbol{\phi}}^T \mathbf{e}}$$

Summary

Reorder & Partition Updated State Space

$$\mathcal{S} = G \cup \overline{G}$$

$G = \{\text{New} + \text{Most affected}\}$ $\overline{G} = \{\text{Less affected}\}$

Summary

Reorder & Partition Updated State Space

$\mathcal{S} = G \cup \overline{G}$

$G = \{\text{New + Most affected}\}$ $\overline{G} = \{\text{Less affected}\}$

Use Less Affected Components From Old Distribution

$\overline{\phi}^T \leftarrow$ components from ϕ^T corresponding to states in $\overline{G}$

Summary

Reorder & Partition Updated State Space

$\mathcal{S} = G \cup \overline{G}$ $G = \{\text{New + Most affected}\}$ $\overline{G} = \{\text{Less affected}\}$

Use Less Affected Components From Old Distribution

$\overline{\phi}^T \leftarrow$ components from ϕ^T corresponding to states in $\overline{G}$

Approximate Censored Distribution

$$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$$

Summary

Reorder & Partition Updated State Space

$\mathcal{S} = G \cup \overline{G}$ $\quad G = \{\text{New + Most affected}\} \quad \overline{G} = \{\text{Less affected}\}$

Use Less Affected Components From Old Distribution

$\overline{\phi}^T \leftarrow$ components from ϕ^T corresponding to states in $\overline{G}$

Approximate Censored Distribution

$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$

Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Summary

Reorder & Partition Updated State Space

$\mathcal{S} = G \cup \overline{G}$ $\qquad G = \{\text{New + Most affected}\} \quad \overline{G} = \{\text{Less affected}\}$

Use Less Affected Components From Old Distribution

$\overline{\phi}^T \leftarrow$ components from ϕ^T corresponding to states in $\overline{G}$

Approximate Censored Distribution

$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$

Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Compute Approximate Aggregated Distribution

$$\boldsymbol{\alpha}^T \leftarrow (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

Summary

Reorder & Partition Updated State Space

$\mathcal{S} = G \cup \overline{G}$ $\quad G = \{\text{New + Most affected}\} \quad \overline{G} = \{\text{Less affected}\}$

Use Less Affected Components From Old Distribution

$\overline{\phi}^T \leftarrow$ components from ϕ^T corresponding to states in $\overline{G}$

Approximate Censored Distribution

$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$

Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Compute Approximate Aggregated Distribution

$\alpha^T \leftarrow (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$

Approximate Updated Distribution

$\boldsymbol{\pi}^T \leftarrow (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}^T)$

Summary

Reorder & Partition Updated State Space

$\mathcal{S} = G \cup \overline{G}$ $\qquad G = \{\text{New} + \text{Most affected}\} \quad \overline{G} = \{\text{Less affected}\}$

Use Less Affected Components From Old Distribution

$\overline{\phi}^T \leftarrow$ components from ϕ^T corresponding to states in $\overline{G}$

Approximate Censored Distribution

$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$

Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Compute Approximate Aggregated Distribution

$\alpha^T \leftarrow (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$

Approximate Updated Distribution

$\pi^T \leftarrow (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}^T)$

Iterate ?

Summary

Reorder & Partition Updated State Space

$\mathcal{S} = G \cup \overline{G}$ $\qquad G = \{\text{New + Most affected}\} \quad \overline{G} = \{\text{Less affected}\}$

Use Less Affected Components From Old Distribution

$\overline{\phi}^T \leftarrow$ components from ϕ^T corresponding to states in $\overline{G}$

Approximate Censored Distribution

$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$

Form Approximate Aggregation Matrix

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

Compute Approximate Aggregated Distribution

$\alpha^T \leftarrow (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$

Approximate Updated Distribution

$\pi^T \leftarrow (\alpha_1, \ldots, \alpha_g, \alpha_{g+1}\mathbf{s}^T)$

Iterate ? **No! — At A Fixed Point**

Iterative Aggregation

Move Off Of Fixed Point With Power Step

$$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$$

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

$$\boldsymbol{\alpha}^T \leftarrow \left(\alpha_1, \ldots, \alpha_g, \alpha_{g+1}\right)$$

$$\boldsymbol{\pi}^T \leftarrow (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}^T)$$

$$\boldsymbol{\psi}^T \leftarrow \boldsymbol{\pi}^T\mathbf{P}$$

(Also makes progress toward convergence when aperiodic)

If $\|\psi^T - \chi^T\| < \tau$ then quit — else

$$\mathbf{s}^T \longleftarrow \overline{\psi^T} / (\overline{\psi^T}\mathbf{e})$$

Iterative Aggregation

Move Off Of Fixed Point With Power Step

$$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$$

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

$$\alpha^T \leftarrow (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

$$\pi^T \leftarrow (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}^T)$$

$$\psi^T \leftarrow \pi^T\mathbf{P}$$

(Also makes progress toward convergence when aperiodic)

If $\|\psi^T - \chi^T\| < \tau$ then quit — else

$$\mathbf{s}^T \longleftarrow \overline{\psi^T} / (\overline{\psi^T}\mathbf{e})$$

Theorem

If $\mathbf{C} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$ is aperiodic, then convergent for all partitions $\mathcal{S} = G \cup \overline{G}$

Iterative Aggregation

Move Off Of Fixed Point With Power Step

$$\mathbf{s}^T \leftarrow \overline{\phi}^T / (\overline{\phi}^T \mathbf{e})$$

$$\mathcal{A} \leftarrow \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12}\mathbf{e} \\ \mathbf{s}^T\mathbf{P}_{21} & 1 - \mathbf{s}^T\mathbf{P}_{21}\mathbf{e} \end{bmatrix}_{(g+1)\times(g+1)}$$

$$\alpha^T \leftarrow (\alpha_1, \ldots, \alpha_g, \alpha_{g+1})$$

$$\pi^T \leftarrow (\alpha_1, \ldots, \alpha_g, \mid \alpha_{g+1}\mathbf{s}^T)$$

$$\psi^T \leftarrow \pi^T \mathbf{P}$$

(Also makes progress toward convergence when aperiodic)

If $\|\psi^T - \chi^T\| < \tau$ then quit — else

$$\mathbf{s}^T \longleftarrow \overline{\psi^T} / (\overline{\psi^T}\mathbf{e})$$

Theorem

If $\mathbf{C} = \mathbf{P}_{22} + \mathbf{P}_{21}(\mathbf{I} - \mathbf{P}_{11})^{-1}\mathbf{P}_{12}$ is aperiodic, then convergent for all partitions $\mathcal{S} = G \cup \overline{G}$

— $|\lambda_2(\mathbf{C})|$ determines rate of convergence

Google's PageRank

Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} \mathbf{1}/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \rightarrow \mathcal{P}_j, \\ \mathbf{0} & \text{otherwise} \end{cases}$$

Google's PageRank

Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} 1/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \rightarrow \mathcal{P}_j, \\ 0 & \text{otherwise} \end{cases}$$

Google Matrix

$$\mathbf{P} = \alpha(\mathbf{H} + \mathbf{E}) + (1 - \alpha)\mathbf{F}$$

- $(\mathbf{H} + \mathbf{E})$ & $\mathbf{F}$ are stochastic $\quad rank(\mathbf{E}) = rank(\mathbf{F}) = 1$
- $0 < \alpha < 1$

Google's PageRank

Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} 1/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \to \mathcal{P}_j, \\ 0 & \text{otherwise} \end{cases}$$

Google Matrix

$$\mathbf{P} = \alpha(\mathbf{H} + \mathbf{E}) + (1 - \alpha)\mathbf{F}$$

— $(\mathbf{H} + \mathbf{E})$ & $\mathbf{F}$ are stochastic $\quad rank(\mathbf{E}) = rank(\mathbf{F}) = 1$

— $0 < \alpha < 1$

— PageRank = $\boldsymbol{\pi}^T$

Google's PageRank

Random Walk On WWW Link Structure

$$\mathbf{H}_{ij} = \begin{cases} 1/(\text{total \# outlinks from page } \mathcal{P}_i) & \text{if } \mathcal{P}_i \to \mathcal{P}_j, \\ 0 & \text{otherwise} \end{cases}$$

Google Matrix

$$\mathbf{P} = \alpha(\mathbf{H} + \mathbf{E}) + (1 - \alpha)\mathbf{F}$$

— $(\mathbf{H} + \mathbf{E})$ & $\mathbf{F}$ are stochastic $\quad rank(\mathbf{E}) = rank(\mathbf{F}) = 1$

— $0 < \alpha < 1$

— PageRank = $\boldsymbol{\pi}^T$

Power Law Distribution

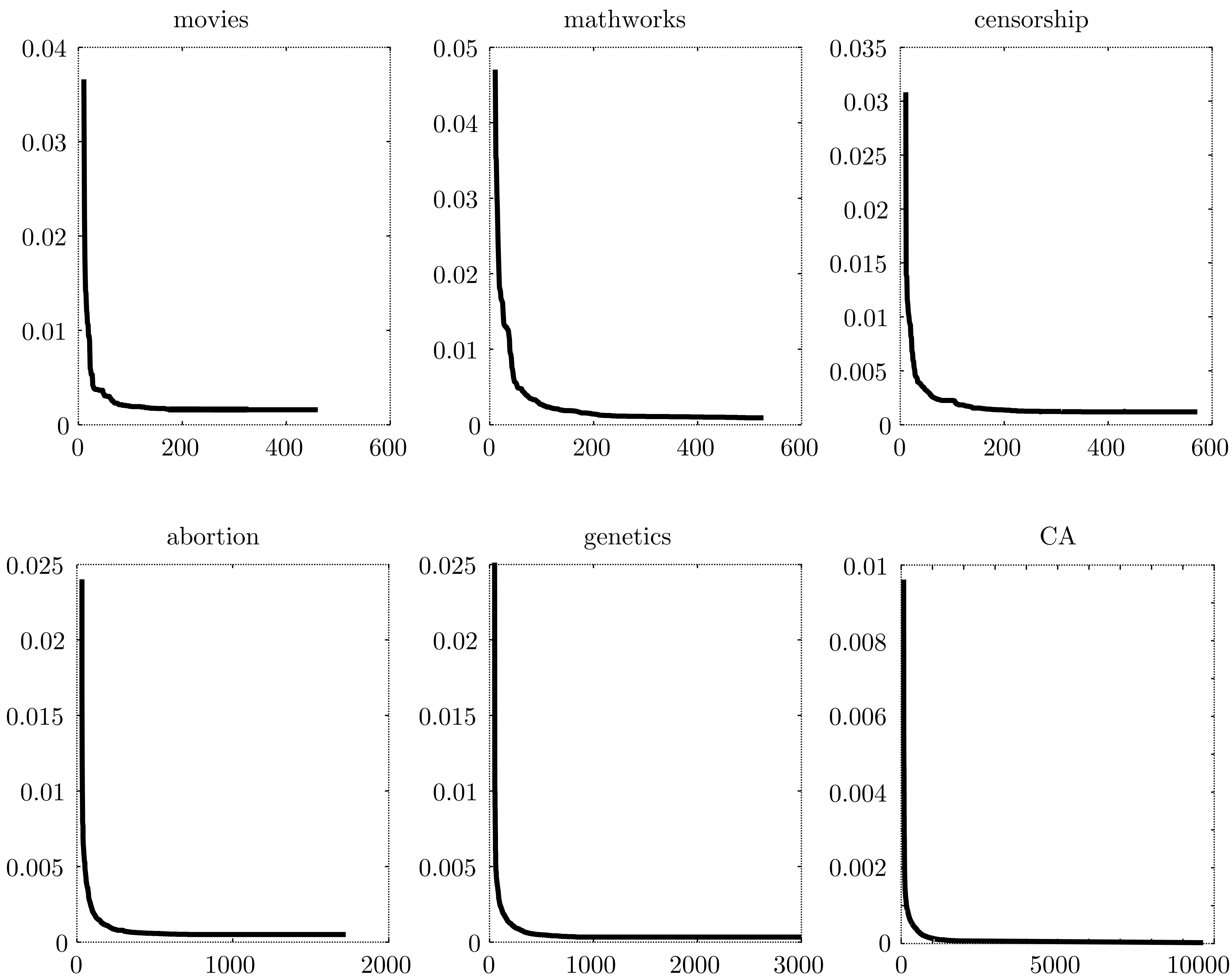
If ordered by magnitude $\pi(1) \geq \pi(2) \geq \cdots \geq \pi(n)$, then

— $\pi(i) \approx \alpha i^{-k}$ for $k \approx \mathbf{2.109}$

[Donato, Laura, Leonardi, 2002]
[Pandurangan, Raghavan,& Upfal, 2004]

— Relatively few large states (i.e., important sites)

"L" Curves



movies
0.04
0.03
0.02
0.01
0
0
200
400
600
mathworks
0.05
0.04
0.03
0.02
0.01
0
0
200
400
600
censorship
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
200
400
600
abortion
0.025
0.02
0.015
0.01
0.005
0
0
1000
2000
genetics
0.025
0.02
0.015
0.01
0.005
0
0
1000
2000
3000
CA
0.01
0.008
0.006
0.004
0.002
0
0
5000
10000

Experiments

The Updates

Nodes Added = 3

Nodes Removed = 50

Links Added = 10 (Different values have little effect on results)

Links Removed = 20

Stopping Criterion

1-norm of residual $< 10^{-10}$

Movies

Power Method

Iterations	Time
17	**.40**

Iterative Aggregation

$\lvert G \rvert$	Iterations	Time
5	**12**	**.39**
10	**12**	**.37**
15	**11**	**.36**
20	**11**	**.35**
25	**11**	**.31**
50	**9**	**.31**
100	**9**	**.33**
200	**8**	**.35**
300	**7**	**.39**
400	**6**	**.47**

$nodes = \mathbf{451}$ $links = \mathbf{713}$

Censorship

Power Method

Iterations	Time
38	1.40

Iterative Aggregation

$\lvert G \rvert$	Iterations	Time
5	38	1.68
10	38	1.66
15	38	1.56
20	20	1.06
25	20	1.05
50	10	.69
100	8	.55
200	6	.53
300	6	.65
400	5	.70

$nodes$ = **562** $links$ = **736**

MathWorks

Power Method

Iterations	Time
54	1.25

Iterative Aggregation

$\|G\|$	Iterations	Time
5	53	1.18
10	52	1.29
15	52	1.23
20	42	1.05
25	20	1.13
50	18	.70
100	16	.70
200	13	.70
300	11	.83
400	10	1.01

$nodes = \mathbf{517} \quad links = \mathbf{13,531}$

Abortion

Power Method

Iterations	Time
106	**37.08**

Iterative Aggregation

$\|G\|$	Iterations	Time
5	**109**	**38.56**
10	**105**	**36.02**
15	**107**	**38.05**
20	**107**	**38.45**
25	**97**	**34.81**
50	**53**	**18.80**
100	**13**	**5.18**
250	**12**	**5.62**
500	**6**	**5.21**
750	**5**	**10.22**
1000	**5**	**14.61**

$nodes = \mathbf{1,693} \quad links = \mathbf{4,325}$

Genetics

Power Method

Iterations	Time
92	**91.78**

Iterative Aggregation

$\lvert G \rvert$	Iterations	Time
5	**91**	**88.22**
10	**92**	**92.12**
20	**71**	**72.53**
50	**25**	**25.42**
100	**19**	**20.72**
250	**13**	**14.97**
500	**7**	**11.14**
1000	**5**	**17.76**
1500	**5**	**31.84**

$nodes = \mathbf{2,952} \quad links = \mathbf{6,485}$

California

Power Method

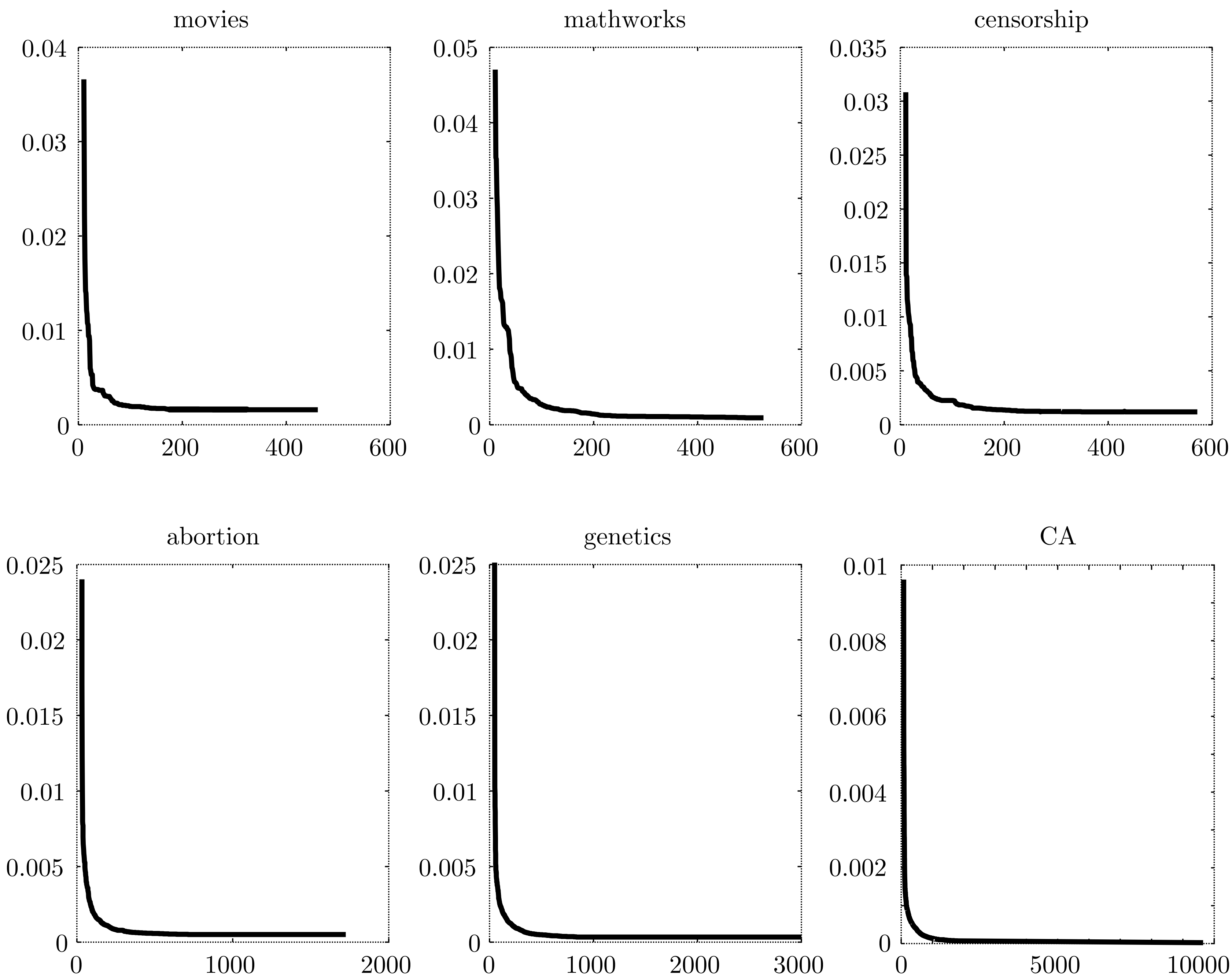
Iterations	Time
176	**5.85**

Iterative Aggregation

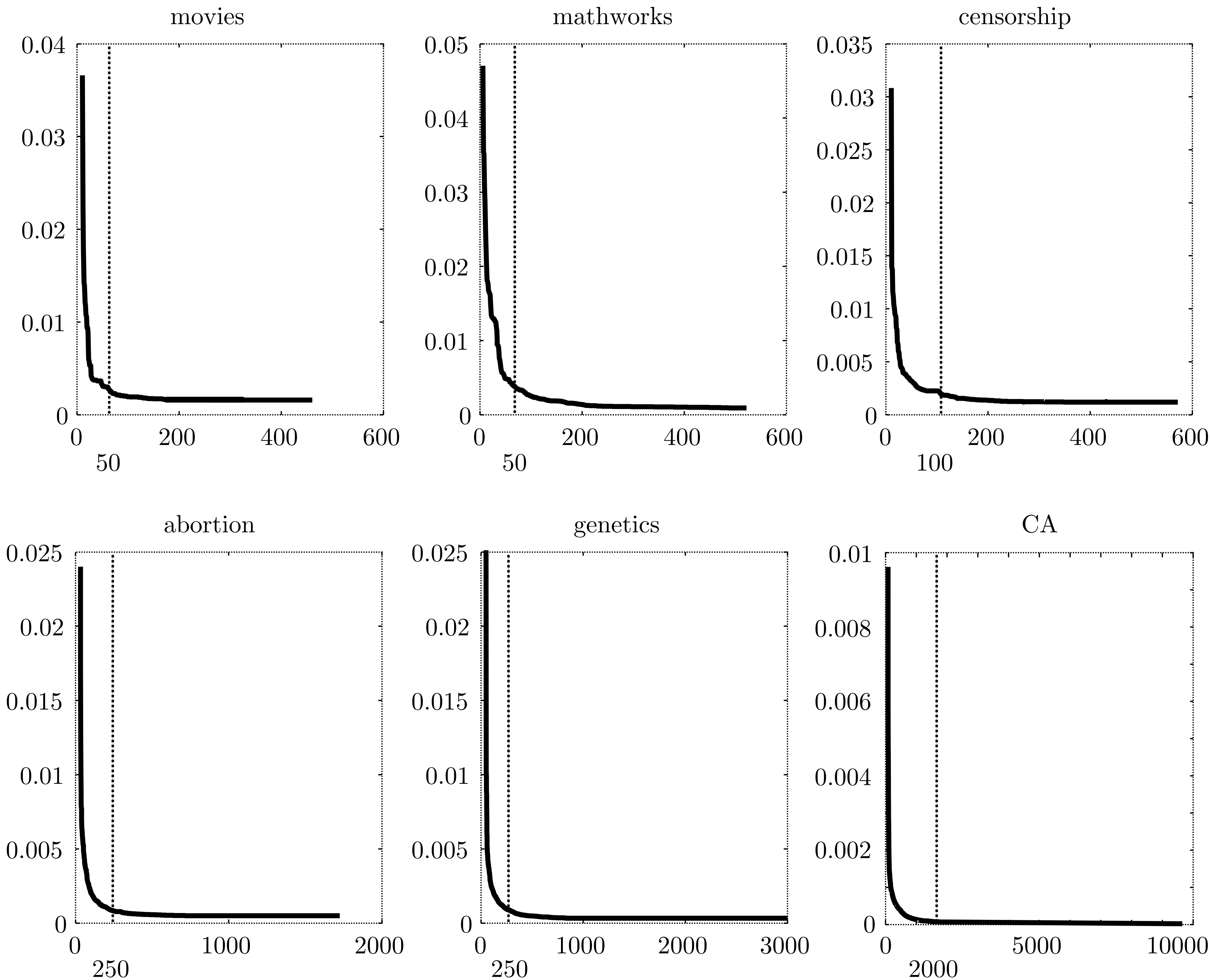
$\|G\|$	Iterations	Time
500	**19**	**1.12**
1000	**15**	**.92**
1250	**20**	**1.04**
1500	**14**	**.90**
2000	**13**	**1.17**
5000	**6**	**1.25**

$nodes = \mathbf{9,664} \quad links = \mathbf{16,150}$

"L" Curves


movies
mathworks
censorship
abortion
genetics
CA
0.04
0.03
0.02
0.01
0
200
400
600
0.05
0.035
0.025
0.015
0.005
1000
2000
3000
0.008
0.006
0.004
0.002
5000
10000

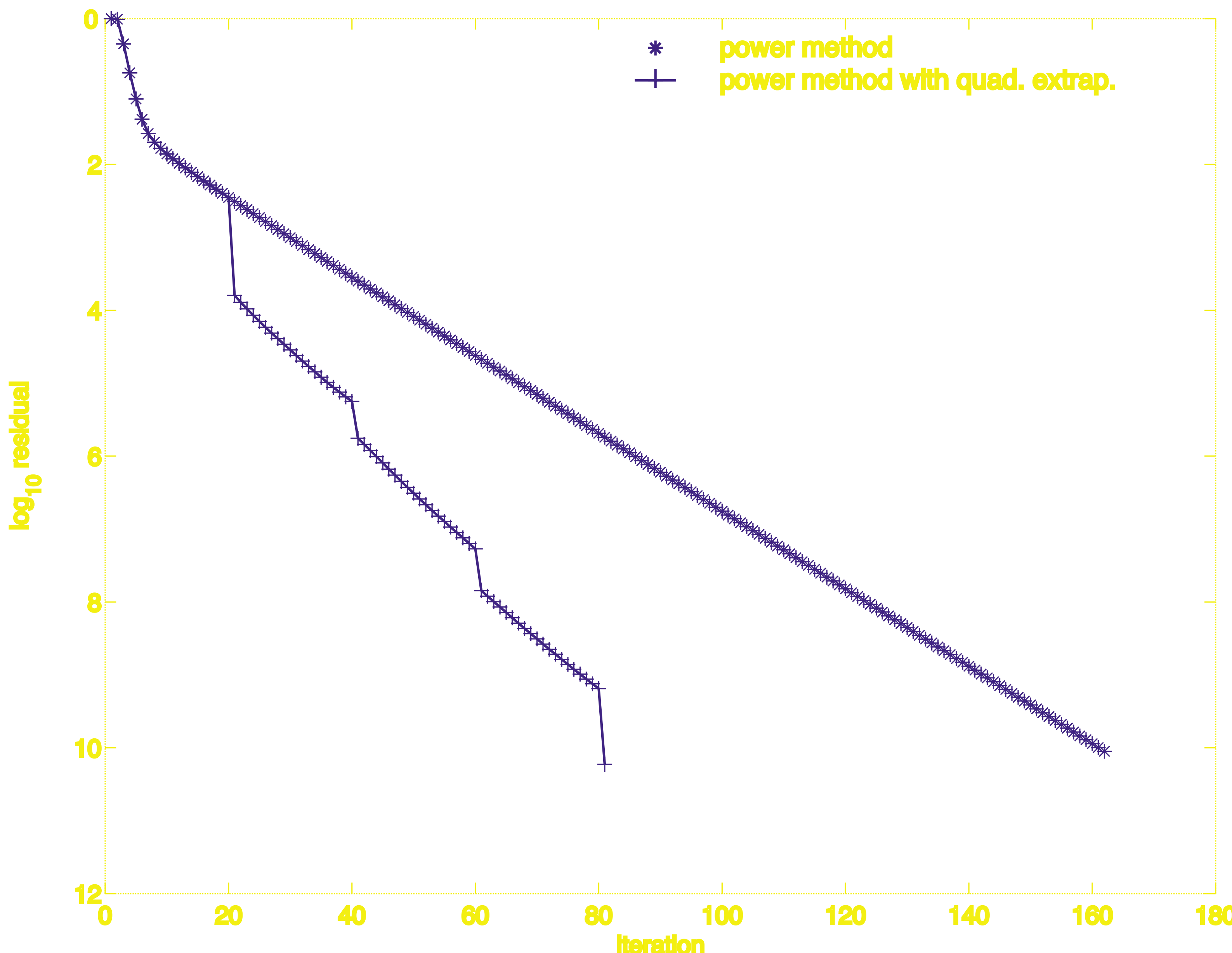
“L” Curves


movies
0.04
0.03
0.02
0.01
0
0
200
400
600
50
mathworks
0.05
0.04
0.03
0.02
0.01
0
0
200
400
600
50
censorship
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
200
400
600
100
abortion
0.025
0.02
0.015
0.01
0.005
0
0
1000
2000
250
genetics
0.025
0.02
0.015
0.01
0.005
0
0
1000
2000
3000
250
CA
0.01
0.008
0.006
0.004
0.002
0
0
5000
10000
2000

Quadratic Extrapolation

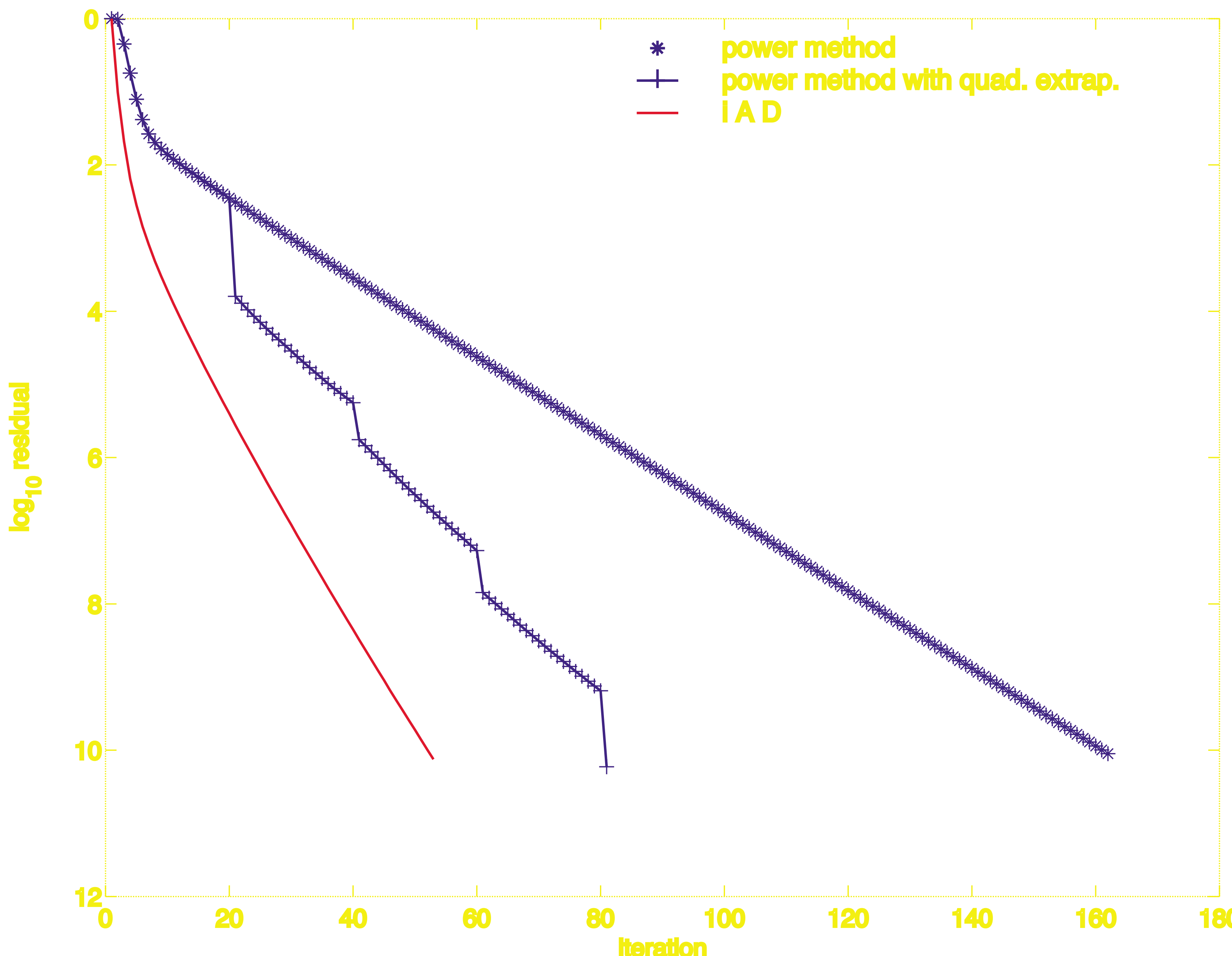
nodes = 10,000 links = 101,118

[Kamvar, Haveliwala, Manning, Golub, 2003]



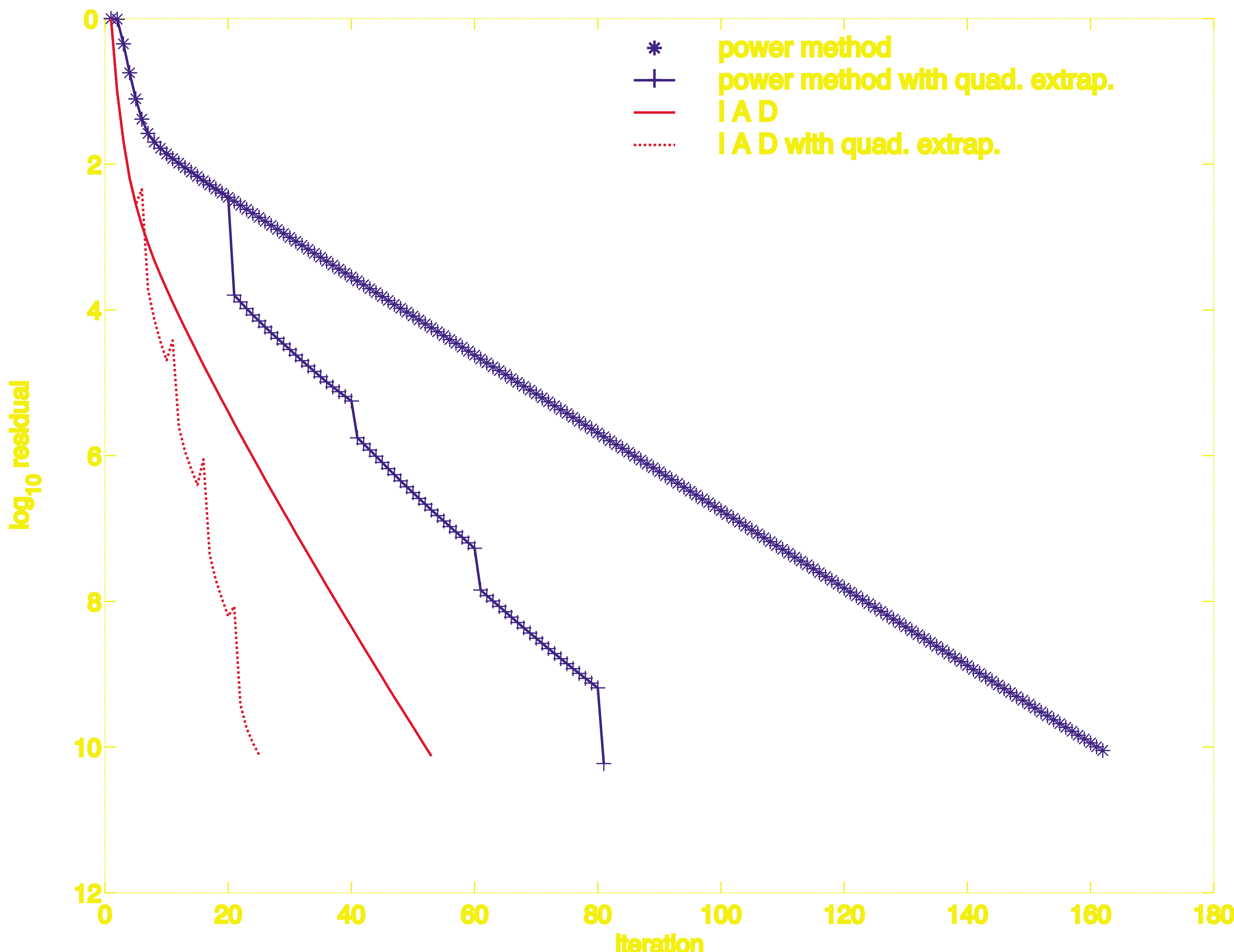
Quadratic Extrapolation

nodes = 10,000   links = 101,118   [Kamvar, Haveliwala, Manning, Golub, 2003]



Quadratic Extrapolation

nodes = 10,000 links = 101,118 [Kamvar, Haveliwala, Manning, Golub, 2003]



Timings

	Iterations	Time (sec)	$\lvert G \rvert$
Power	162	9.69	
Power+Quad	81	5.93	
IAD	21	2.22	2000
IAD+Quad	16	1.85	2000

$nodes = 10,000 \quad links = 101,118$

Conclusion

- Iterative aggregation shows promise for updating Markov chains
- Especially for those having power law distributions

Leveling Off Point

$$\pi(i) \approx \alpha i^{-k}$$

$$\left|\frac{d\pi(i)}{di}\right| \approx \epsilon$$ for some user-defined tolerance ϵ

$$i_{level} \approx \left(\frac{k\alpha}{\epsilon}\right)^{\mathbf{1/}k\mathbf{+1}}$$

Perhaps better: $$i_{level} \approx f(n) \left(\frac{k\alpha}{\epsilon}\right)^{\mathbf{1/}k\mathbf{+1}}$$

For WWW: $$g_{opt} \approx f(n) \left[\frac{\mathbf{2.109}\alpha}{\epsilon}\right]^{\mathbf{1/3.109}}$$