A Preconditioned Power Method for Computing Stationary Vectors of Markov Chains, with Application to Internet Search Engines

Ilse Ipsen

Steve Kirkland

Overview

- Simple Web Model
- Power Method
- Exploiting Structure
- ILU Preconditioned Power Method
- Google Matrix

Simple Web Model

Determine importance of a web page

PageRank of page:
Probability that surfer visits page

Page i has d outgoing links

If page i has no link to page j then $p_{ij} = 0$
else $p_{ij} = 1/d$

P is stochastic matrix

$$0 \le p_{ij} \le 1 \qquad P\mathbf{1} = \mathbf{1}$$

Eigenvectors of P

$P\mathbf{1} = \mathbf{1} \;\Rightarrow\; P$ has eigenvalue 1

left eigenvector: $\pi^T P = \pi^T,\ \pi \geq 0,\ \|\pi\|_1 = 1$

ith entry of π: probability that surfer visits page i
PageRank of page i

PageRank $\doteq$ largest left eigenvector of P

Power Method

Initial vector $v_{(0)} > 0$, $\|v_{(0)}\|_1 = 1$
After k iterations: $v_{(k)}^T \leftarrow v_{(0)}^T P^k$

If P is primitive then:

$\pi^T P = \pi^T$ with $\pi > 0$, $\|\pi\|_1 = 1$, unique
Power method converges: $v_{(k)} \rightarrow \pi$
Convergence rate $|\boldsymbol{\lambda_2}| < \mathbf{1}$

Primitive Matrices

Matrix $P \geq 0$ is primitive if $P^m > 0$ for some $m \geq 1$

P is irreducible: $P \ncong \begin{pmatrix} X & X \\ 0 & X \end{pmatrix}$

If P stochastic then eigenvalue 1 is distinct

$\rho(P)$ produced by single eigenvalue

If P stochastic then eigenvalues $\neq 1$ have magnitude < 1

$(P^m > 0 \;\Rightarrow\; \rho(P^m)$ simple)

Power Method

Initial vector $v_{(0)} > 0$, $\|v_{(0)}\|_1 = 1$

After k iterations: $v_{(k)}^T \leftarrow v_{(0)}^T P^k$

If P is primitive then:

$\pi^T P = \pi^T$ with $\pi > 0$, $\|\pi\|_1 = 1$, unique

Power method converges: $v_{(k)} \rightarrow \pi$

Convergence rate $|\boldsymbol{\lambda_2}| < \mathbf{1}$

Slow convergence if $|\lambda_2| \approx 1$

Can we exploit the stochastic structure of P?

Exploiting Structure in P

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \qquad I - P_{11} \quad \text{nonsingular}$$

Factor $I - P = LDU$

$$L = \begin{pmatrix} I & \\ * & I \end{pmatrix} \; U = \begin{pmatrix} I & * \\ & I \end{pmatrix} \; D = \begin{pmatrix} I - P_{11} & \\ & I - S \end{pmatrix}$$

where

$$S = P_{22} + P_{21}(I - P_{11})^{-1}P_{12}$$

stochastic complement

LDU in Eigenvalue Problem

Use $I - P = LDU$:

$$\pi^T P = \pi^T \quad \Leftrightarrow \quad \pi^T(I - P) = 0 \quad \Leftrightarrow \quad \pi^T LD = 0$$

$$\begin{pmatrix} \pi_1^T & \pi_2^T \end{pmatrix} \begin{pmatrix} I - P_{11} & \\ -P_{21} & I - S \end{pmatrix} = 0$$

$$\pi_2^T(I - S) = 0 \qquad \pi_1^T = \pi_2^T \, P_{21}(I - P_{11})^{-1}$$

π_2 is eigenvector for smaller matrix S

Smaller Eigenvalue Problem

$$S = P_{22} + P_{21}(I - P_{11})^{-1}P_{12}$$

If P is stochastic and irreducible then

- S is stochastic and irreducible
- $\sigma^T S = \sigma^T$ with $\sigma > 0$, $\|\sigma\|_1 = 1$, unique

If S is primitive then

- Power method converges
- Convergence rate $|\boldsymbol{\lambda_2}(\mathbf{S})| < \mathbf{1}$

Idea

- Exact: $\pi_2^T = \pi_2^T S$

$$\pi^T = \left(\pi_2^T P_{21}(I - P_{11})^{-1} \quad \pi_2^T S\right)$$

- Approximation: Any $\tilde{\pi}_2 > 0$

$$\tilde{\pi}^T = \left(\tilde{\pi}_2^T P_{21}(I - P_{11})^{-1} \quad \tilde{\pi}_2^T S\right)$$

- Repeat with: $\tilde{\pi}_2 := \tilde{\pi}_2^T S$

The Big Picture

$$\tilde{\pi}^T = \begin{pmatrix} \tilde{\pi}_2^T P_{21}(I - P_{11})^{-1} & \tilde{\pi}_2^T S \end{pmatrix}$$

$$= \begin{pmatrix} * & \tilde{\pi}_2^T \end{pmatrix} \begin{pmatrix} 0 & 0 \\ P_{21}(I - P_{11})^{-1} & S \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 \\ P_{21}(I - P_{11})^{-1} & S \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} L^{-1} P$$

ILU Preconditioned Power method ($I - P = LDU$)

ILU Preconditioned Power Method

- Initial guess: $\pi_{(0)} > 0, \quad \|\pi_{(0)}\| = 1$
- Iterate: $\pi^T_{(k+1)} = \rho_k \, \pi^T_{(k)} \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} L^{-1} P$

  $$\|\pi_{(k+1)}\| = 1$$

$$\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} L^{-1} P = \begin{pmatrix} 0 & 0 \\ P_{21}(I - P_{11})^{-1} & S \end{pmatrix}$$

Convergence rate $|\lambda_2(S)|$

But . . .

- ILU preconditioned power method requires formation of $S = P_{22} + P_{21}(I - P_{11})^{-1}P_{12}$ in

$$\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} L^{-1}P = \begin{pmatrix} 0 & 0 \\ * & S \end{pmatrix}$$

- Does ILU preconditioned power method converge faster than power method?

Is $|\lambda_2(S)| < |\lambda_2(P)|$?

How To Avoid Forming S

$$\tilde{\pi}^T = \rho \begin{pmatrix} * & \tau^T \end{pmatrix} \begin{pmatrix} 0 & 0 \\ P_{21}(I - P_{11})^{-1} & S \end{pmatrix} \qquad \|\tilde{\pi}\|_1 = 1$$

is mathematically equivalent to:

Form $A = \begin{pmatrix} P_{11} & P_{12}\mathbf{1} \\ \tau^T P_{21} & \tau^T P_{22}\mathbf{1} \end{pmatrix}$

Compute $\alpha > 0$ where $\alpha^T A = \alpha^T$, $\|\alpha\|_1 = 1$

Partition $\alpha = \begin{pmatrix} \omega^T & \rho \end{pmatrix}$

Multiply $\tilde{\pi}^T = \begin{pmatrix} \omega^T & \rho\,\tau^T \end{pmatrix} P$

Connections

ILU Preconditioned power method $\doteq$
Iterative Aggregation/Disaggregation

- IAD based on stochastic complementation:
  Meyer 1989, Meyer 2004 (SIAM)
  Langville & Meyer 2002, 2003, 2004
  Billy Stewart 1985, 1992, 1994
  Lee, Golub & Zenios 2003
- IAD based on general splittings:
  Marek & Szyld 1994, Szyld 2004 (ILAS)
  Marek & Mayer 1998, 2001, 2003
- Haveliwala, Kamvar, Klein, Manning & Golub 2003
  Kamvar, Haveliwala & Golub 2004
  Choi & Saunders (SIAM 2004)

Upper Bounds for $|\lambda_2|$

- If P stochastic then $|\lambda_2(P)| \le \tau(P)$

  where $$\tau(P) \equiv \frac{1}{2} \max_{i,j} \|e_i^T P - e_j^T P\|_1$$

- If P also irreducible then $\tau(S) \le \tau(P)$

  where $$S = P_{22} + P_{21}(I - P_{11})^{-1} P_{12}$$

Upper bound for $|\lambda_2(S)| \le$ upper bound for $|\lambda_2(P)|$

Google Matrix

$$G = cP + (1-c)\mathbf{1}v^T$$

P is stochastic matrix

$0 < c < 1$, $v > 0$ and $\|v\|_1 = 1$

- Power method for G always converges:

$$|\lambda_2(G)| \leq c < 1$$

- ILU preconditioned power method can converge faster than power method:

  If $|\lambda_2(P)| = 1$ then $|\lambda_2(S_G)| \leq |\lambda_2(G)|$

Assumptions

$$G = cP + (1-c)\mathbf{1}v^T$$

where P is stochastic such that

- P contains essential index classes $C_1, \ldots, C_k$
- Each C_j contains index i_j such that $P_{i_j, i_j} = 0$

$$QGQ^T = \begin{array}{cc} & k \\ k & \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \end{array} \quad Q \text{ permutation}$$

where G_{11} contains rows & columns $i_1 \ldots i_k$ of G

ILU Power Method: Google Matrix

$$S_G \equiv G_{22} + G_{21}(I - G_{11})^{-1} G_{12}$$

Under previous assumptions:

$$|\lambda_2(S_G)| < |\lambda_2(G)|$$

With appropriate partitioning:

ILU preconditioned power method converges faster than power method

Summary

- ILU preconditioned power method
- Power method on stochastic complement S
- Convergence rate $|\lambda_2(S)|$
- Upper bounds for $|\lambda_2(S)|$
- Implemented as:
  Iterative Aggregation/Disaggregation method
- No need to form S
- Faster convergence for Google matrix than power method