Detecting Overlapping Communities by finding Sparse Vectors in Unconverged Invariant Subspace Approximations

David Bindel

Department of Computer Science
Cornell University

18 Apr 2012
I have no idea what you're talking about...

...so here's a bunny with a pancake on its head.
Why eigenvalues?
A play in three acts.

David Bindel

Department of Computer Science
Cornell University

18 Apr 2012
A complicated PDE

\[ \rho \frac{\partial^2 u}{\partial t^2} - E \frac{\partial^2 u}{\partial x^2} = 0 \]

Produces a simple solution

\[ u(x, t) = u_0(x) \sin(\omega t) \]
Resonating MEMS and Tiny Guitars

Microguitars from Cornell University (1997 and 2003)

- MEMS = Micro-Electro-Mechanical Systems
- Micron-scale *mechanical* structures with IC fab technology
- Widely used for sensing and signal processing ...
- ... and sometimes really high-pitch guitars!
Modeling a Disk Resonator
Modeling a Disk Resonator

\[ V_{\text{in}} \]

\[ C_0 \]

\[ L_x \]

\[ C_x \]

\[ R_x \]

\[ V_{\text{out}} \]
Modeling a Ringing Disk

- At what frequencies does this vibrate?
- How quickly is the ringing damped?
- What about errors (in numerics or fabrication)?
- How do we answer these questions \textit{fast}?
Current example: Micro-HRG / GOBLiT / OMG

- This is a gyroscope!
- Now make it 1mm across.
- Collaborator roles:
  - Basic design
  - Fabrication
  - Measurement
- Our part:
  - Detailed physics
  - Fast software
  - Sensitivity
  - Design optimization
A little GOBLiT
Outline

1. Act I: Physics
2. Act II: Graphs
3. Act III: Games
4. Concluding thoughts
Graph Bisection

Goal: Cut in half, minimize edges cut.
Graphs to Quadratics

Give node $i$ a label $x_i = \pm 1$.

- The labels cut the graph in half:

$$\sum_{i=1}^{n} x_i = 0.$$  

- Count cut edges by a quadratic form

$$|\text{cut edges}| = \frac{1}{4} \sum_{(i,j) \in E} (x_i - x_j)^2$$
Quadratic forms and matrices

\[
|\text{cut edges}| = \frac{1}{4} \sum_{(i,j)\in E} (x_i - x_j)^2
\]

\[
= \frac{1}{4} \sum_{(i,j)\in E} \begin{bmatrix} x_i \\ x_j \end{bmatrix}^T \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ x_j \end{bmatrix} = \frac{1}{4} x^T L x
\]

where

\[
L_{ij} = \begin{cases} \text{degree of node } i, & i = j \\ -1, & (i,j) \in E \end{cases}
\]
Graph bisection

Idea: Minimize $s^T L s$ such that $e^T s = 0$, $s \in \{\pm 1\}^n$.

Oops — NP hard!
Relax!

Hard: \( \min \ s^T L s \) \ s.t. \( e^T s = 0, \ s \in \{ \pm 1 \}^n \).

Easy: \( \min \ v^T L v \) \ s.t. \( e^T v = 0, \ v \in \mathbb{R}^n, \|v\|^2 = n \).

This is an eigenvalue problem!
Three cups, three straws, a drop of dye
The random walker

Basic idea: Ideas diffuse, too!

Day 1: I came up with a funny joke!
Day 2: I tell everyone in my family
Day 3: My mother tells a friend?

Ideas diffuse fastest within communities (graph clusters).
Simon-Ando theory

Markov chain with loosely-coupled subchains:

- Rapid *local* mixing: after a few steps

\[ p_k \approx \sum_{j=1}^{c} \alpha_{j,k} p^{(j)}_{\infty} \]

where \( p^{(j)}_{\infty} \) is a local equilibrium for the \( j \)th subchain

- Slow equilibration: \( \alpha_{j,k} \to \alpha_{j,\infty} \).
Spectral Simon-Ando picture

Exactly decoupled case ($c$ decoupled chains):
- Eigenvalue one has multiplicity $c$.
- Eigenvectors of $T$ are local equilibria.
- Rapid mixing $\implies$ large gap to $\lambda_{c+1}$.

Weakly coupled case:
- Cluster of $c$ eigenvalues near 1.
- Eigenvectors of $T$ are combinations of local equilibria.
- Large gap between $\lambda_c$ and $\lambda_{c+1}$.
Outline

1. Act I: Physics
2. Act II: Graphs
3. Act III: Games
4. Concluding thoughts
How Bad is Choosing Your Own Opinion?

Alice ➔ Paul ➔ Bob
Carol ➔ Paul ➔ David
Modeling Opinion Formation

A basic model:

- A fixed *intrinsic* opinion $s_i$
- A variable *expressed* opinion $x_i$
- Equilibrium $x_i = \text{argmin}_{z_i} c_i(z_i)$, where

$$c_i(z_i) \equiv (s_i - z_i)^2 + \sum_{j \in N(i)} w_{ij}(z_i - x_j)^2$$

- Define a *social cost* $c(z) = \sum_i c_i(z_i)$
Methodology: Graph problem $\mapsto$ linear algebra problem.

Nash equilibrium: $$(L + I)x = s$$

Social optimum: $$(A + I)y = s$$

Cost at equilibrium: $$c(x) = s^T Cs$$

Optimal social cost: $$c(y) = s^T Bs$$

Price of anarchy is a ratio of quadratics:

$$\text{PoA}(s) = \frac{c(x)}{c(y)} = \frac{s^T Cs}{s^T Bs}$$

Find worst case through a generalized eigenvalue problem:

$$Cs_* = \lambda Bs_*$$
Applications abound!
Why eigenvalue analysis?

- Because it simplifies many problems.
- It’s a nonlinear equation I can solve.
- It’s a nonconvex optimization I can solve.

... and because I’ve been thinking about it for a while!
Why scientific computing?

Because connections are fun!