Algorithms for Evolving Networks

Des Higham
Department of Mathematics and Statistics
University of Strathclyde

SCMS/DFG Sep. 2010
D.F. Griffiths, University of Dundee, UK; D.J. Higham, University of Strathclyde, Glasgow, UK

**Numerical Methods for Ordinary Differential Equations**

**Initial Value Problems**

Numerical Methods for Ordinary Differential Equations is a self-contained introduction to a fundamental field of numerical analysis and scientific computation. Written for undergraduate students with a mathematical background, this book focuses on the analysis of numerical methods without losing sight of the practical nature of the subject. It covers the topics traditionally treated in a first course, but also highlights new and emerging themes. Chapters are broken down into ‘lecture’ sized pieces, motivated and illustrated by numerous theoretical and computational examples. Over 200 exercises are provided and these are starred according to their degree of difficulty. Solutions to all exercises are available to authorized instructors. The book covers key foundation topics:

- Taylor series methods
- Runge--Kutta methods
- Linear multistep methods
- Convergence
- Stability
- Adaptive stepsize selection
- Long term dynamics
- Modified equations
- Geometric integration
- Stochastic differential equations

The prerequisite of a basic university-level calculus class is assumed, although appropriate background results are also summarized in appendices. A dedicated website for the book... more on [http://springer.com/978-0-85729-147-9](http://springer.com/978-0-85729-147-9)

- Focuses on the analysis of numerical methods without losing sight of the practical nature of the subject
- Covers topics traditionally treated in a first course, but also highlights new and emerging themes
- Features chapters broken down into lecture-sized pieces, motivated and illustrated by numerous theoretical and computational examples

---

Order Now!

Yes, please send me ____ copies

"Numerical Methods for Ordinary Differential Equations"
Some previous work:
- communicability
- bipartite communities
- directed bipartite communities

Recent extension:
- evolving networks

We take a linear algebra viewpoint and derive spectral methods.
We also quantify statistical significance.
Collaborators

- **Strathclyde**: Jonathan Crofts, Ernesto Estrada, Martin McDonald, Chanpen Phokaew, Clare Lee, Alan Taylor, Xiaolin Xiao
- **Dundee**: Keith Vass
- **Bath**: Alastair Spence, Zhivkho Stoyanov
- **Reading**: Peter Grindrod, Mark Parsons
- **Oxford**: Tim Behrens, Heidi Johansen-Berg
- **Beatson Labs**: Gabriela Kalna
- **Imperial**: Natasa Przulj
Unweighted, undirected, with $N$ nodes.

Adjacency matrix $A$.

$$(A^2)_{ij} := \sum_{p=1}^{N} a_{ip}a_{pj}$$

counts the number of paths of length two from node $i$ to node $j$.

Generally, $$(A^k)_{ij}$$ counts the number of walks of length $k$
from node $i$ to node $j$.

Communicability between distinct nodes $i$ and $j$:

$$\left( A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots + \frac{A^k}{k!} + \cdots \right)_{ij}, \text{ that is, } (e^A)_{ij}.$$
9 subjects - at least 6 months following first, left-hemisphere, subcortical stroke; and 10 age matched controls.

Diffusion Tensor Imaging computes all connections between all voxels

Connectivity network based on the Harvard-Oxford cortical and subcortical structural atlas: 48 cortical regions and 8 subcortical regions.

Methodology from FMRIB Oxford: Heidi Johanson-Berg, Tim Behrens, Saad Jbabdi.
Stroke data supplied by Rose Bosnell.
New weighted communicability measure gives a statistically significant improvement.
Connections within the RHS only:

Crofts, Higham, Bosnell, Jbabdi, Matthews, Behrens, Johansen-Berg,

**NeuroImage**, to appear

Evidence for compensatory rewiring of RHS for stroke victims.
Drill down to brain regions:
Bipartivity in Protein-Protein Interaction


Yellow: share **SH3** domain. Brown: part of the **actin cortical patch assembly mechanism of vesicle endocytosis** (Drees et al. 2001).

Separate experiments (Kessels & Qualman 2004, Freisen et al. 2005) show **SH3** involved in **trafficking of vesicles**.

[vesicle: small, enclosed compartment within a cell.]
Distinct subsets of nodes $S_1$ and $S_2$ such that

- $S_1$ has few internal links
- $S_2$ has few internal links
- there are many $S_1 \rightarrow S_2$ links
- few other links involve $S_1$ or $S_2$
An alternating walk of length $k$ from node $i_1$ to node $i_{k+1}$
is a list of nodes

$$i_1, i_2, i_3, \ldots, i_{k+1}$$
such that $a_{i_s, i_{s+1}} \neq 0$ for $s$ odd, and $a_{i_{s+1}, i_s} \neq 0$ for $s$ even.

Loosely, an alternating walk is a traversal that successively follows links in the forward and reverse directions.
This motivates…

\[ f(A) = I - A + \frac{AA^T}{2!} - \frac{AA^T A}{3!} + \frac{AA^T AA^T}{4!} - \cdots \]

We expect \( f(A)_{ij} \) to take

- large positive values when \( i, j \in S_1 \), and
- large negative values when \( i \in S_2 \) and \( j \in S_1 \).

In terms of the SVD, \( A = U \Sigma V^T \), we have

\[ f(A) = U \cosh(\Sigma) U^T - U \sinh(\Sigma) V^T \]
Analogously, similar comments apply to $f(A^T)$.

**Overall idea:** $f(A) + f(A^T)$ has

- positive values representing inter-community $S_1 \leftrightarrow S_1$ and $S_2 \leftrightarrow S_2$ relationships, and
- negative values representing extra-community $S_1 \leftrightarrow S_2$ relationships

Also, $f(A) + f(A^T)$ is a symmetric matrix, so amenable to standard clustering techniques.
Example

\[ A \]

\[ \exp(A) \]

\[ \exp(-A) \]

\[ f(A) + f(A^T) \]
Definition for a walk and its length? 

Note the lack of symmetry caused by time's arrow.
Definition for a walk and its length?

Note the lack of symmetry caused by time's arrow.
Definition for a walk and its length?

Note the lack of symmetry caused by time’s arrow.
Definition for a **walk** and its **length**?  

Note the **lack of symmetry** caused by time’s arrow.
Time points $t_0 < t_1 < t_2 < \cdots < t_M$

Adjacency matrices $A^{[0]}, A^{[1]}, A^{[2]}, \ldots, A^{[M]}$

Applications include
- communication
- social or on-line interaction
- neuroscience

Modelling issues and epidemic spread considered in:

**Dynamic walk of length** $w$ **from node** $i_1$ **to node** $i_{w+1}$:
sequence of times $t_{r_1} \leq t_{r_2} \leq \cdots \leq t_{r_w}$ and a
sequence of edges $i_1 \leftrightarrow i_2, i_2 \leftrightarrow i_3, \ldots, i_w \leftrightarrow i_{w+1}$,
such that $i_m \leftrightarrow i_{m+1}$ exists at time $t_{r_m}$
Key observation: the matrix product

\[ A^{[r_1]} A^{[r_2]} \ldots A^{[r_w]} \]

has \( i, j \) element that counts the number of dynamic walks of length \( w \) from node \( i \) to node \( j \), where the \( m \)th step takes place at time \( t_{r_m} \).

Keep track of all such walks and discount by \( \alpha^w \)

Use the expansion \( (I - \alpha A)^{-1} = I + \alpha A + \alpha^2 A^2 + \cdots \)

This motivates the algorithm:
Set \( B^{[M]} = (I - \alpha A^{[M]})^{-1} \) and let

\[ B^{[s]} = (I - \alpha A^{[s]})^{-1} B^{[s+1]}, \quad \text{for} \ s = M - 1, M - 2, \ldots, 0 \]

Then \( (B^{[0]})_{ij} \) is our overall summary of how well information can be passed from node \( i \) to node \( j \).
We will call the row and column sums

$$\sum_{k=1}^{N} (B[0])_{nk} \quad \& \quad \sum_{k=1}^{N} (B[0])_{kn}$$

the **broadcast** and **receive** communicabilities

- generalizes **eigenvector centrality** in social networks
- involves **sparse linear solves**
- captures the **asymmetry** through non-commutativity of matrix multiplication

Related work by Tang, Musolesi, Mascolo, Latora & Nicosia, *3rd Workshop on Social Network Systems, 2010* uses shortest paths and considers different issues.
Synthetic Test: $N = 1001$, $M = 30$, $\alpha = 0.2$

Sequence of independent Erdös/Rényi/Gilbert random graphs on 1000 nodes, with average degree 2. At each time node 1001 was connected to the two nodes with largest degree.
Weekly voice call contact at MIT

Broadcast (source) and receive (sink) communicability
151 nodes over 1138 consecutive days
(many days are empty)

\[
\text{Log(Total Degree)} \quad \text{Log(Receive)} \\
\text{Log(Total Degree)} \quad \text{Log(Broadcast)}
\]

\[
\ln(\text{degree}) \ vs \ \ln(\text{broadcast}) \quad \ln(\text{degree}) \ vs \ \ln(\text{receive})
\]
Top 15 Nodes

Total Degree
48, 73, 67, 17, 9, 107, 137, 49, 114, 20, 70, 100, 76, 147, 18

Broadcast
114, 70, 72, 32, 55, 141, 20, 19, 2, 28, 4, 30, 67, 18, 3

Receive
48, 67, 50, 96, 13, 147, 107, 75, 137, 49, 122, 88, 40, 76, 136
<table>
<thead>
<tr>
<th>Degree</th>
<th>Broadcast</th>
<th>Receive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tana Jones</td>
<td>Chris Germany</td>
<td>Tana Jones</td>
</tr>
<tr>
<td>Jeff Dasovich</td>
<td>Michelle Lokay</td>
<td>Sara Shackleton</td>
</tr>
<tr>
<td>Sara Shackleton</td>
<td>Mark McConnell</td>
<td>Susan Bailey</td>
</tr>
<tr>
<td>Mike Grisby</td>
<td>Shelley Corman</td>
<td>Scott Neal</td>
</tr>
<tr>
<td>Stephanie Panus</td>
<td>Paul Barbo</td>
<td>Marie Heard</td>
</tr>
<tr>
<td>Louse Kitchen</td>
<td>Steven Harris</td>
<td>Carol Clair</td>
</tr>
<tr>
<td>Drew Fossum</td>
<td>Kimberly Watson</td>
<td>Louise Kitchen</td>
</tr>
<tr>
<td>Susan Scott</td>
<td>Lindy Donoho</td>
<td>Gerald Nemec</td>
</tr>
<tr>
<td>Chris Germany</td>
<td>Kevin Hyatt</td>
<td>Drew Fossum</td>
</tr>
<tr>
<td>Kimberly Watson</td>
<td>Darrell Schoolcraft</td>
<td>Susan Scott</td>
</tr>
</tbody>
</table>
Linear algebra based on counting walks gives a natural approach for finding patterns in networks and summarizing key features.

The idea extends readily to evolving networks, capturing the inherent asymmetry, and leading to dynamic centrality measures.

Current issues include:
- Sensitivity to time resolution
- Optimal downweighting strategy
- Dealing with weighted edges