Learning Directed Acyclic Graphs from Group-Structure Data

Yongsu Lee

Department of Statistics
University of Wisconsin-Madison

Sep 4, 2020
Usually, for learning DAG ...

What if ... ?

Group 1

Group 2

Group 3
Zheng et al. (2018) suggests NOTEARS method ...

\[
E[X_1|X_2, X_3] = \beta_{01} + \beta_{21}X_2 + \beta_{31}X_3 \\
E[X_2|X_1, X_3] = \beta_{02} + \beta_{12}X_1 + \beta_{32}X_3 \\
E[X_3|X_1, X_2] = \beta_{03} + \beta_{13}X_1 + \beta_{23}X_2
\]

\[
W = \begin{bmatrix}
\beta_{12} & \beta_{13} \\
\beta_{21} & \beta_{23} \\
\beta_{31} & \beta_{32}
\end{bmatrix}
\]

Loss function

Lasso Penalized term

\[
\arg\min_W \ell(W) + \lambda \|W\|_1
\]

Elementwise product

subject to \( \text{tr}(e^{W^\top W}) - p = 0 \)  
\underline{Acyclicity restriction}
$\mathbf{W} =$

\[
\begin{bmatrix}
\|\mathbf{W}^{(1,1)}\|_F & \|\mathbf{W}^{(1,2)}\|_F & \|\mathbf{W}^{(1,3)}\|_F \\
\|\mathbf{W}^{(2,1)}\|_F & \|\mathbf{W}^{(2,2)}\|_F & \|\mathbf{W}^{(2,3)}\|_F \\
\|\mathbf{W}^{(3,1)}\|_F & \|\mathbf{W}^{(3,2)}\|_F & \|\mathbf{W}^{(3,3)}\|_F
\end{bmatrix}
\]

Index matrices with proper dimensions
Newly proposed NOTEARGIS method is ....

\[
\arg\min_W \ell(W) + \lambda \sum_{i \neq j} ||W_{(i,j)}||_F
\]

subject to \( \text{tr}(e^{A(W \circ W)B}) - d = 0 \)

acyclicity restriction among groups

---

Why NOTEARGIS ?

- Easily applicable to any type of variables
- Can be extended to a mixed-type DAG network learning.
- More precise result for structure learning.
**Grouped Gaussian Data**
10 groups (nodes) – 2 elements for each 50 observations; 50 simulation iterations

**Multi-level Data**
30 nodes – 2, 3, or 4 levels 50 observations; 50 simulation iterations

<table>
<thead>
<tr>
<th>Method</th>
<th>E</th>
<th>TPR</th>
<th>FDR</th>
<th>SHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOTEARGIS</td>
<td>7.04</td>
<td>0.7</td>
<td>0.3</td>
<td>5.34</td>
</tr>
<tr>
<td>NOTEARS</td>
<td>3.06</td>
<td>0.31</td>
<td>0.4</td>
<td>8.48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>E</th>
<th>TPR</th>
<th>FDR</th>
<th>SHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOTEARGIS</td>
<td>7.88</td>
<td>0.26</td>
<td>0.68</td>
<td>45.52</td>
</tr>
<tr>
<td>CD</td>
<td>4.46</td>
<td>0.15</td>
<td>0.9</td>
<td>70.94</td>
</tr>
</tbody>
</table>