Variational Inference of Sparse Network from Count Data

Julien Chiquet, Mahendra Mariadasou, Stéphane Robin

AgroParisTech/INRA (French National Institute for Agricultural Research)

ICML, Long Beach Convention Center, June the 11th 2019

R/C++ PLNmodels package, development version on github

```
install.packages("PLNmodels")
```

https://jchiquet.github.io/PLNmodels/

Sparse multivariate Poisson Lognormal model

A sparse latent multivariate Gaussian model

$$\mathbf{Z}_i \text{ iid} \sim \mathcal{N}_p(\mathbf{0}_p, \mathbf{\Omega}^{-1}), \qquad \mathbf{\Omega} \text{ sparse}, \quad \|\mathbf{\Omega}\|_{1,\text{offdiagonal}} < c$$

$$\mathbf{Y}_i \,|\, \mathbf{Z}_i \sim \mathcal{P}(\exp\{\mathbf{O}_i + \mathbf{X}_i^\top \mathbf{B} + \mathbf{Z}_i\}) \qquad (i,j) \notin \mathcal{E} \Leftrightarrow Z_i \perp\!\!\!\perp Z_j | Z_{\setminus\{i,j\}} \Leftrightarrow \mathbf{\Omega}_{ij} = 0.$$

Interpretation

- Dependency structure (network) encoded in the latent space ($\mathbf{\Omega}$)
- Additional effects due to covariates $\mathbf{X}$ are fixed
- Conditional Poisson distribution = noise model

Sparse Variational Inference

Variational approximation

Take $q_i \equiv \mathcal{N}(\mathbf{m}_i, \text{diag}(\mathbf{s}_i))$ to approximate of $p(Z_i | Y_i)$

model parameters $\boldsymbol{\theta} = (\mathbf{B}, \boldsymbol{\Omega})$

variational parameters $\boldsymbol{\psi} = (\mathbf{M}, \mathbf{S})$ where $\mathbf{M} = [\mathbf{m}_1^\top \dots \mathbf{m}_n^\top]^\top$, $\mathbf{S} = [(\mathbf{s}_1^2)^\top \dots (\mathbf{s}_n^2)^\top]^\top$

Sparse lower bound of the likelihood

$$J(\boldsymbol{\theta}, \boldsymbol{\psi}) - \lambda \left\| \boldsymbol{\Omega} \right\|_{1,\text{off}} = \mathbb{E}_q[\log p_{\boldsymbol{\theta}}(\mathbf{Y}, \mathbf{Z})] + \mathcal{H}[q_\psi(\mathbf{Z})] - \lambda \left\| \boldsymbol{\Omega} \right\|_{1,\text{off}}.$$

Alternate optimization – objective is biconcave in $(\mathbf{B}, \mathbf{M}, \mathbf{S})$ and $\boldsymbol{\Omega}$

1. $(\hat{\mathbf{B}}, \hat{\mathbf{M}}, \hat{\mathbf{S}})$: gradient ascent
2. $\hat{\boldsymbol{\Omega}}$: graphical-Lasso problem

Selection of λ – StARS (*Stability Approach to Regularization Selection*)

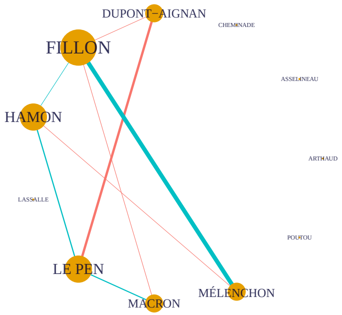
Illustration: first round of French Presidential 2017

source: https://data.gouv.fr

- data: votes cast for each of the 11 candidates in the more than 63,000 polling stations
- offset: log-registered population of voter (account for different station sizes)
- covariate: "department" (administrative division, a proxy for geography)

⇝ find *competing* candidates, who appeal to different voters, and *compatible* candidates

Inferred network of partial correlation (blue: negative, red: positive)



Latent Positions (PCA)