UC DAVIS
UNIVERSITY OF CALIFORNIA

Network Regression with Graph Laplacians

Yidong Zhou, Hans-Georg Müller

Department of Statistics, University of California, Davis

NEURAL INFORMATION PROCESSING SYSTEMS

Motivation

- **Samples of networks** find applications in diverse fields such as traffic mobility and brain connectivity.
- Addressing the statistical challenges posed by samples of networks is an emerging frontier, where **each network is treated as an individual data point**.
- The network space presents a significant challenge due to **the absence of a linear structure**, a crucial feature in traditional statistical methodologies.
- Drawing inspiration from foundational parametric and nonparametric regression models designed for scalar responses, particularly **linear and local linear regression**, becomes a compelling avenue for addressing the complexities associated with network responses.

Characterization of the Network Space

$G=(V,E)$: a network with a set of nodes $V=\{v_1,\dots,v_m\}$ and a set of edge weights $E=\{w_{ij}: w_{ij}\geq 0, i,j=1,\dots,m\}$, where $w_{ij}=0$ indicates v_i and v_j are unconnected.

(C0) G_m is **simple**, i.e., there are no self-loops or multi-edges.
(C1) G_m is **weighted**, **undirected**, and **labeled**.
(C2) Edge weights w_{ij} are bounded above by $W\geq 0$, i.e., $0\leq w_{ij}\leq W$.

Any network satisfying Conditions (C0)–(C2) can be uniquely associated with its graph Laplacian $L=(l_{ij})$, defined as

$$l_{ij}=\begin{cases}-w_{ij}, & i\neq j\\ \sum_{k\neq i} w_{ik}, & i=j\end{cases}$$

for $i,j=1,\dots,m$. The network space can thus be characterized by the corresponding space of graph Laplacians,

$$\mathcal{L}_m=\{L=(l_{ij}): L=L';\ L1_m=0_m;\ -W\leq l_{ij}\leq 0 \text{ for } i\neq j\}, \qquad (1)$$

where 1_m and 0_m are the m-vectors of ones and zeroes, respectively.

Proposition 1

The space $\mathcal{L}_m$, defined in (1), is a ***bounded***, ***closed***, *and* ***convex*** *subset in $\mathbb{R}^{m^2}$ of dimension $m(m-1)/2$.*

The convexity and closedness of $\mathcal{L}_m$, as demonstrated in Proposition 1, ensure the **existence** and **uniqueness** of projections onto $\mathcal{L}_m$, which will be utilized in the proposed regression approach.

Choice of Metrics

Although the space of graph Laplacians $\mathcal{L}_m$ lacks a linear structure, it is generally considered a metric space when equipped with an appropriate metric. There are various metrics to choose from for $\mathcal{L}_m$ and one common choice is the **Frobenius metric**, defined as

$$d_F(L_1,L_2)=[\mathrm{tr}\{(L_1-L_2)'(L_1-L_2)\}]^{1/2}.$$

While d_F is the simplest among the possible metrics on $\mathcal{L}_m$, it exhibits a **swelling effect**, particularly for positive definite matrices.

Let $\mathcal{S}_m^+$ denote the space of real symmetric positive semi-definite $m\times m$ matrices. Another popular metric, designed to mitigate the swelling effect, is the **power metric**,

$$d_{F,\alpha}(L_1,L_2)=d_F\{F_\alpha(L_1),F_\alpha(L2)\},$$

where $F_\alpha(S)=U\Lambda^\alpha U'$, and $U\Lambda U'$ represents the spectral decomposition of $S\in\mathcal{S}_m^+$. For $\alpha=1$, $d_{F,\alpha}$ reduces to the Frobenius metric d_F. For $\alpha=1/2$, the square root metric $d_{F,1/2}$ is a canonical choice that has been widely studied.

Fréchet Mean and Conditional Fréchet Mean

Consider the random pair $(X,L)\sim F$, where X takes values in $\mathbb{R}^p$, $L\in\mathcal{L}_m$ is a graph Laplacian, and F denotes a suitable probability law. We investigate the dependence of L on covariates of interest X.

The Fréchet mean and conditional Fréchet mean are generalizations of the mean and conditional mean from the real line to general metric spaces. Let $Y\in\mathbb{R}$ denote a random variable on the real line and let d represent either the Frobenius or power metric.

- Note that $E(Y)=\operatorname*{arg\,min}_{y\in\mathbb{R}} E\{(Y-y)^2\}$ and $E(Y|X)=\operatorname*{arg\,min}_{y\in\mathbb{R}} E\{(Y-y)^2|X\}$.
- Mean $\longrightarrow$ Fréchet mean:
$$E(Y)\longrightarrow \operatorname*{arg\,min}_{\omega\in\mathcal{L}_m} E\{d^2(L,\omega)\}.$$
- Conditional mean $\longrightarrow$ conditional Fréchet mean:
$$E(Y|X)\longrightarrow \operatorname*{arg\,min}_{\omega\in\mathcal{L}_m} E\{d^2(L,\omega)|X\}.$$

Global and Local Network Regression

To model the relationship between networks and Euclidean predictors, a natural target is the conditional Fréchet mean. Recall that for scalar responses, linear regression assumes a linear relationship between X and the conditional mean of Y given X, i.e.,

$$E(Y|X)=\beta_0+\beta_1'X,$$

where β_1 denotes the slope vector. Using ordinary least squares, the regression function can be alternatively characterized by

$$E(Y|X=x)=E\{s_G(x)Y\}=\operatorname*{arg\,min}_{y\in\mathbb{R}} E\{s_G(x)(Y-y)^2\},$$

where the weight function $s_G(x)=1+(X-\mu)'\Sigma^{-1}(x-\mu)$ with $\mu=E(X)$ and $\Sigma=\mathrm{Var}(X)$.

The **global network regression**, extending linear regression to network responses, is defined as

$$m_G(x)=\operatorname*{arg\,min}_{\omega\in\mathcal{L}_m} E\{s_G(x)d^2(L,\omega)\}, \qquad (2)$$

where d can be either the Frobenius or power metric. Suppose that $(X_k,L_k)\sim F$, $k=1,\dots,n$ are independent and define

$$\bar{X}=\frac{1}{n}\sum_{k=1}^n X_k,\quad \hat{\Sigma}=\frac{1}{n}\sum_{k=1}^n (X_k-\bar{X})(X_k-\bar{X})'$$

as the sample mean and covariance. The regression function in (2) can be estimated by

$$\hat{m}_G(x)=\operatorname*{arg\,min}_{\omega\in\mathcal{L}_m}\frac{1}{n}\sum_{k=1}^n s_{kG}(x)d^2(L_k,\omega), \qquad (3)$$

where $s_{kG}(x)=1+(X_k-\bar{X})'\hat{\Sigma}^{-1}(x-\bar{X})$.

The **local network regression**, a generalization of local linear regression to network responses, follows a similar form but employs a different weight function.

Theorem 1

Let the space of graph Laplacians $\mathcal{L}_m$ be endowed with the Frobenius metric d_F. Then for a fixed $x\in\mathbb{R}^p$, it holds for $m_G(x)$ and $\hat{m}_G(x)$ as per (2) and (3) that,

$$d_F\{m_G(x),\hat{m}_G(x)\}=O_p(n^{-1/2}).$$

Furthermore, for a given $B>0$ and any $\varepsilon>0$,

$$\sup_{\|x\|\leq B} d_F\{m_G(x),\hat{m}_G(x)\}=O_p(n^{-1/\{2(1+\varepsilon)\}}).$$

Remark

- *The global network regression estimate in (3) using the Frobenius metric d_F can be simplified as a* ***projection*** *$P_{\mathcal{L}_m}$ onto $\mathcal{L}_m$,*
$$\hat{m}_G(x)=\operatorname*{arg\,min}_{\omega\in\mathcal{L}_m} d_F^2\{B_G(x),\omega\}=P_{\mathcal{L}_m}\{B_G(x)\},\quad B_G(x)=\frac{1}{n}\sum_{k=1}^n s_{kG}(x)L_k.$$
- *Both* ***pointwise and uniform rates of convergence*** *have been established for global and local network regression, utilizing both the Frobenius and power metrics.*
- *Pointwise rates of convergence are* ***optimal*** *for both global and local network regression when employing the Frobenius and power metrics with $0<\alpha<1$.*

New York Yellow Taxi System After COVID-19 Outbreak

- **Responses**: traffic networks from Apr 12 to Sep 30, 2020 in Manhattan.
- **Predictors**: COVID-19 new cases per day, weekend indicator.
- **Model**: global network regression.

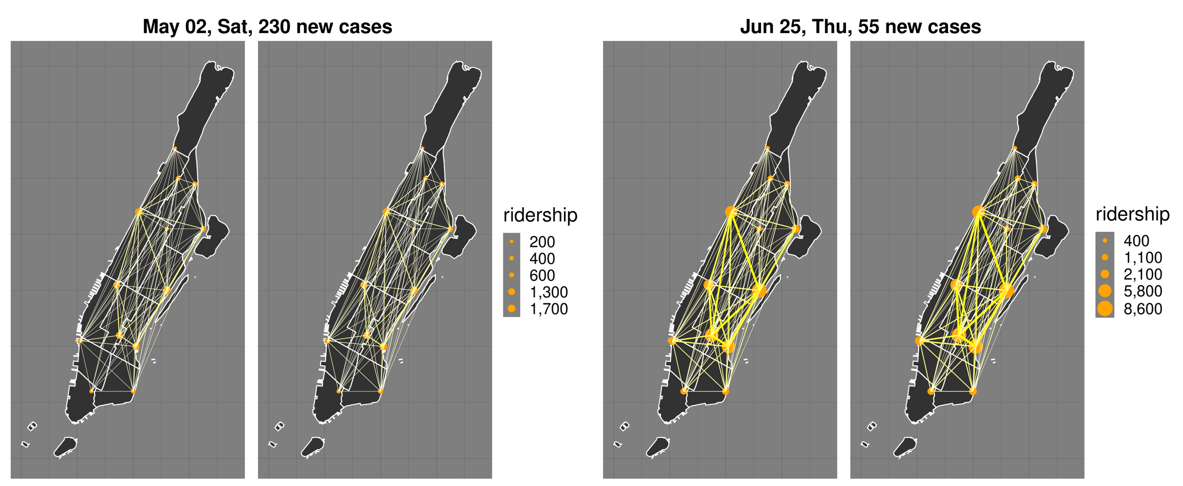

Figure 1. True (left) and fitted (right) networks on May 2, 2020 and Jun 25, 2020. The corresponding number of COVID-19 new cases are in the headline.

Dynamics of Networks in the Aging Brain

- **Responses**: functional brain connectivity networks.
- **Predictors**: age.
- **Model**: local network regression.
- The number of communities for ages 65, 70, 75, and 80 is 10, 12, 12, and 16, respectively.
- These communities are found to be associated with different anatomical regions of the brain.
- Higher age is associated with increased local interconnectivity and cliquishness.

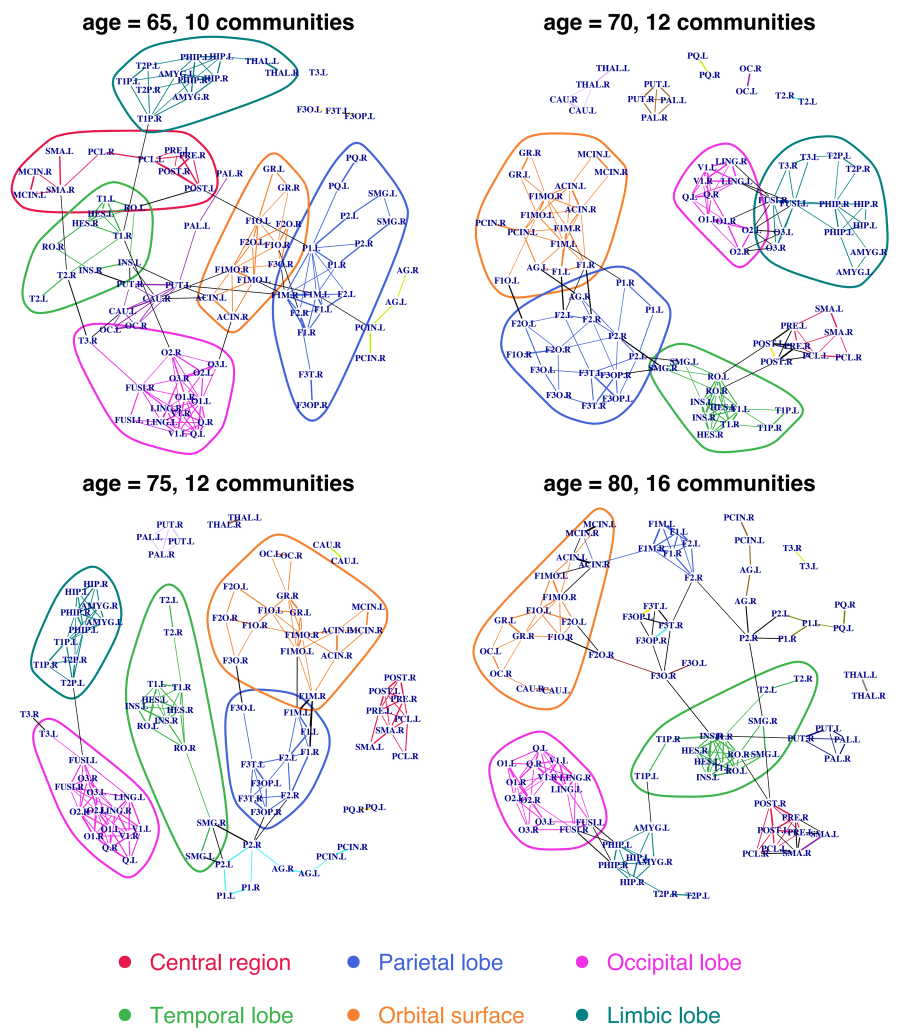

Figure 2. Topological representation using spectral community detection for predicted functional connectivity networks at different ages (years). The communities comprising 10 or more regions of interest are highlighted using colored polygons.

Relevant Literature

- Zhou, Y., & Müller, H. G. (2022). Network regression with graph Laplacians. *Journal of Machine Learning Research*, 23(320), 1-41.
- Petersen, A., & Müller, H. G. (2019). Fréchet regression for random objects with Euclidean predictors. *The Annals of Statistics*, 47(2), 691-719.
- Chen, Y., & Müller, H. G. (2022). Uniform convergence of local Fréchet regression with applications to locating extrema and time warping for metric space valued trajectories. *The Annals of Statistics*, 50(3), 1573-1592.