Index

May 19, 2020

\#
Generalized Modeling and Predictions in Multiscale Geographically Weighted Regression

---

0.0.1 Proposal Abstract:

A recent addition to the local statistical models in PySAL is the implementation of Multiscale Geographically Weighted Regression (MGWR) model, a multiscale extension to the widely used approach for modeling process spatial heterogeneity - Geographically Weighted Regression (GWR). GWR is a local spatial multivariate statistical modeling technique embedded within the regression framework that is calibrated and estimates covariate parameters at each location using borrowed data from neighboring observations. The extent of neighboring observations used for calibration is interpreted as the indicator of scale for the spatial processes and is assumed to be constant across covariates in GWR. MGWR, using a back-fitting algorithm relaxes the assumption that all processes being modeled operate at the same spatial scale and estimates a unique indicator of scale for each process.

The GWR model in PySAL can currently estimate Gaussian, Poisson and Logistic models though the MGWR model is currently limited to only Gaussian models. This project aims to expand the MGWR model to nonlinear local spatial regression modeling techniques where the response outcomes may be discrete (following a Poisson distribution) or binary (Logistic models). This will enable a richer and holistic local statistical modeling framework to model multi-scale process heterogeneity for the open source community. In addition, to support efficient testing for different model implementations, a simulated data generator module will be implemented to supply test datasets following unique model variable distribution needs. This will also provide a foundation for possible expansion to test other local model implementations in PySAL. GWR has been widely used as a tool for spatial prediction and has been known to be informative on the spatial processes generating the data being predicted (Harris et al., 2010). While the GWR implementation in PySAL facilitates the predictions for the dependent variable at unsampled locations, this functionality has not been implemented for MGWR yet. This project aims to also enable the prediction functionality for MGWR and solve for its growing need in the open source community (https://github.com/pysal/mgwr/issues/51). In doing so, open issues around predictions in GWR (for e.g. https://github.com/pysal/mgwr/issues/50) will also be resolved.

0.0.2 Deliverables (Completed and ongoing):

---

1. Expansion of MGWR model to Poisson dependent variables (Completed)

- **Final Pull Request in pysal/mgwr**
- **Introduction and comprehensive tests with real and simulated datasets**

2. Expansion of MGWR model to Binomial dependent variables (Ongoing)

- **Introduction to approaches attempted and comprehensive tests with real and simulated datasets**

3. Predictions in GWR and MGWR (Ongoing)

- **All development commits through the coding period**

0.0.3 Personal Information

Sub organization information: Python Spatial Analysis Library (PySAL)
Mentors: Levi John Wolf, Wei Kang, Taylor Oshan
Student Information:
Name: Mehak Sachdeva
Github username: mehak-sachdeva
Email: msachde1@asu.edu
Mobile: +1-857-206-9743
Time zone: MST – Mountain Standard Time / Mountain Time
GSoC Blog RSS Feed URL: https://blogs.python-gsoc.org/en/mehaksachdevas-blog/

0.0.4 References:

- Harris, P., Fotheringham, A. S., Crespo, R., & Charlton, M. (2010). The Use of Geographically Weighted Regression for Spatial Prediction: An Evaluation of Models Using Simulated Data Sets. Mathematical Geosciences, 42(6), 657–680. https://doi.org/10.1007/s11004-010-9284-7

- Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. Statistical Science, 1(3), 297–310. https://doi.org/10.1214/ss/1177013604

- Nakaya, T., Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2005). Geographically weighted Poisson regression for disease association mapping. Statistics in Medicine, 24(17), 2695–2717. https://doi.org/10.1002/sim.2129