

Inference on the Effects of Observed Features in Latent Space Models for Networks

Zachary Jones

Matthew Denny

Bruce Desmarais

Hanna Wallach

April 19, 2016

Abstract

The latent space model (LSM) for network data is a popular model that combine a generalized linear model with a latent spatial embedding of the network to model ties. It has been used to decrease error in the estimation of and inference regarding the effect of observed covariates. In applications of the LSM, it is assumed that the latent spatial embedding can control for unmeasured confounders that have an effect on the value of edges in a network. We investigate the latent space model's performance on this task via a Monte Carlo simulation wherein we show that in the presence of an unmeasured covariate which can be appropriately modeled using a latent space, estimation and inferential error decreases, however, the difference is not large, even at high levels of multicollinearity between the unmeasured and observed covariates. However, the prediction error of the LSM when an unmeasured covariate is present is substantially lower in most cases. When an omitted covariate is not present, the LSM can introduce bias and increased inferential error above that of a generalized linear model. We investigate solutions to this issue and conclude with recommendations as to the appropriate use of the latent space model for observational studies.

1 Introduction

1.1 Problem

Latent variable inference, generally conceived, present the possibility of representing unmeasured data in statistical models. The latent space model (LSM) Hoff, Raftery and Handcock (2002) for network data is regularly used to estimate the effect of covariates in the presence of latent network structure. However, it is unclear that the introduction of a latent space decreases the expected error for the parameter(s) of interest when aspects of the network structure (e.g., homophily) that are unmeasured are correlated with measured covariates. There are two reasons that using the LSM may lead to increased error. First, the latent configurations inferred may result in a representation of the network wherein a node's position in the latent space is spuriously correlated with

the observed covariates, leading to reduced efficiency due to multicollinearity. Second, if the unobserved (i.e., latent) network structure is truly correlated with the observed covariates, the unobserved structure that can be correlated with the observed variable may be attributed to the observed variable, while the latent space parameters are used to model other sources of variation.

In what follows we identify the conditions under which using the LSM reduces the estimation and inferential error with respect to the effects of observed covariates. We compare the performance of LSM to the generalized linear model (GLM) as a baseline. We develop a simulation design in which we control the degree of collinearity between omitted network structure and observed covariates. We investigate whether the prior can be tuned to avoid efficiency losses in using the LSM. We conclude with recommendations regarding the ideal use of the LSM.

2 Background

2.1 Applications and Development of the Latent Space Model

The LSM has seen use in a variety of fields in which network data is common, particularly the social sciences. This appeal appears to have been driven primarily by the LSM’s usefulness in modeling transitivity and homophily which appear to be ubiquitous in many (social) networks. In political science the LSM and variations on the form developed in Hoff, Raftery and Handcock (2002) have been used to estimate the effect of democracy on the probability of a militarized interstate dispute (Ward, Siverson and Cao 2007), the amount of portfolio investment between states (Cao and Ward 2013), and the effect of multimember districts on the probability of collaboration between state legislators in the United States (Kirkland 2012). Variations of the LSM developed for networks measured over time have been applied to the study of international trade, wherein the effects of various features of trading partners are estimated (Ward, Ahlquist and Rozenas 2013). In ecology the LSM has been used to study the sociality of elephants (Vance, Archie and Moss 2009) and orcas (Fearnbach, Durban, Ellifrit, Waite, Matkin, Lunsford, Peterson, Barlow and Wade 2014), to discover ecological communities (Fletcher, Acevedo, Reichert, Pias and Kitchens 2011; Fletcher Jr, Revell, Reichert, Kitchens, Dixon and Austin 2013), and to study food webs (Chiu and Westveld 2011). In epidemiology the LSM has been used to identify clusters of infected persons for later isolation (Zhang, Wang, Wang and Fang 2015) and to study patterns of interaction amongst physicians (Paul, Keating, Landon and O’Malley 2014). In marketing and business research the LSM has been used to study inter-group trust (Dass and Kumar 2011), optimal bundling and pricing of goods and brands for retailers (Dass and Kumar 2012). In neuroscience the LSM has been proposed as a method for modelling fMRI data (Simpson, Bowman and Laurienti 2013).

Although the latent space model is arguably most useful as an exploratory or predictive model (Hoff, Raftery and Handcock 2002; Shmueli 2010) it has been applied in some cases in a manner that suggests that users of the model assume that the latent space parameters reduce estimation and/or inferential error with respect to the effects of measured explanatory variables. Though it is not the case that all of the aforementioned articles explicitly or implicitly make such a claim (not all of the above applications of the LSM even include covariates), if the LSM does reduce inferential/estimation error under certain conditions, the LSM might find wider use due to the pervasiveness of network data where there remains unmeasured structure which either confounds relationships between measured covariates and edges, in which case the LSM might be applied to reduce bias, or in which unmeasured structure is related to the strength/presence of edges and can be useful in decreasing estimator variance. Additionally, the predictive performance of the LSM has only seen limited evaluation (Hoff, Raftery and Handcock 2002), despite having been used for model selection (Ward, Ahlquist and Rozenas 2013; Fletcher et al. 2011; Fletcher Jr et al. 2013; Chiu and Westveld 2011). Hence, finding the conditions under which the LSM reduces prediction error also may affect future use of the LSM.

The LSM has also seen a substantial amount of subsequent development and extension. The latent space has been represented as a k -dimensional Euclidean space and by latent factors (Hoff, Raftery and Handcock 2002; Hoff 2009). Additional structure has been introduced by adding random effects, which, for example, may involve sender or receiver specific effects for directed networks which capture differential activity rates amongst nodes (Hoff 2003). Within this framework Westveld and Hoff (2011) models dynamic network data by treating the latent space as a stochastic process. Handcock, Raftery and Tantrum (2007) enable the LSM to model clustering that is not representable as homophily (i.e., stochastic equivalence) by combining the LSM with latent cluster models. Hoff (2008) shows that the LSM and latent cluster models are special cases of an “eigenmodel.” That is, an eigendecomposition of a symmetric sociomatrix can be used to represent both latent space and cluster models, but not vice-versa.

2.2 Collinearity in Generalized Linear Models

(Mackinnon and Puterman 1989) and (Weissfeld and Sereika 1991) investigate multicollinearity in generalized linear models and propose diagnostic criteria. Collinearity of the covariates causes the observed information matrix to be ill-conditioned, which produces numerical instability in the iterative estimation of $\hat{\beta}$, the “effects” of covariates included in a regression. (Weissfeld and Sereika 1991), extending the work of (Davis, Hyde, Bangdiwala and Nelson 1986), propose the use of the condition index and the variance decomposition proportion of each condition index, to diagnose the presence and degree of multicollinearity in generalized linear models. The observed information matrix at the maximum likelihood estimate, $\mathcal{I}(\hat{\beta})$ is scaled to have unit length

and singular value decomposition is applied. The resulting right singular values are the eigenvectors and the associated singular values are the square root of the eigenvalues of the scaled observed information matrix since it is real Hermitian. The condition index is the ratio of the largest eigenvalue to each of the eigenvalues, and is at least 1. The square root of the condition number gives the potential magnification of the numerical error in inverting the observed information matrix. Hence our expectation is that as collinearity increases the variance of estimators of the effect of observed covariates will increase regardless of whether this is omitted confounding structure in the data.

3 Research Design

3.1 Priors for the Latent Space Model

3.1.1 Current Implementations

What priors are currently used and why? Hoff, Raftery and Handcock (2002) propose the use of diffuse independent normal priors for the latent positions and regression parameters. They also recognize the fact that isolates' positions are weakly identified in the LSM, and in one application they actually just exclude the isolate.¹ Hoff, Raftery and Handcock (2002) experiment with the use of an exponential prior for the intercept, which they argue sets a lower bound on the probability of a tie between nodes. Of course, the latent space positions could simply drift further apart in order to compensate for the high intercept. The following paragraph on priors appears on page 1097.

We have not discussed in detail the choice of a prior distribution for latent positions in this article. Although simple, the diffuse independent normal priors presented in the examples may not accurately represent prior beliefs about the structure of social networks. More appropriate might be clustered point processes or mixtures of normals with an unknown number of components. Such priors could allow one to incorporate prior information on tendencies for clustering, without specifying cluster membership. This would add another level of hierarchy to the analysis, although the resulting model would be more flexible and perhaps more accurately represent any tendencies of populations to form segregating groups.

(Krivitsky, Handcock, Raftery and Hoff 2009) implement a hierarchical model in which each node's latent space position is drawn from a mixture of G normals. The covariate priors are single Gaussians.

¹Excluding isolates seems fine if the inferential purpose is to estimate positions in the latent space. However, such deletion would introduce bias in estimating covariate effects.

3.1.2 Alternatives to consider

Would regularization/sparsity priors help with identifying the covariate effects? Mitchell and Beauchamp (1988) propose a class of “spike and slab” prior as a class of priors for the coefficients in linear regression. These priors are well-suited to the problem of variable selection for linear regression. The prior is formed as a mixture between a diffuse uniform prior and a point mass at zero. Such a prior would not be appropriate for latent space coordinates since weakly identified coordinates of isolates or separate components would likely exhibit strange behavior at the boundary of the diffuse uniform. We might consider instead mixing, e.g., a point mass at zero with a distribution such as the Cauchy, which is fairly flat in the tails, but is defined on the entire real line and still exhibits a surface that descends from the center of the distribution. This prior structure would build a preference into the model for using the observed covariates, rather than the latent coordinate parameters, to explain tie formation.

What priors could be used to avoid collinearity between observed covariates and latent space? Look to Bayesian Latent Trait models and specifically differential item functioning (DIF) in the IRT literature.

Should priors be calibrated against observed covariates? Using informative priors assures that the latent space coordinates are weakly identified, meaning that at some distance from the origin the increase in the likelihood from moving isolates away from the other nodes, or components away from each other, is offset by decreases in the prior from moving coordinates away from the origin. However, we may also use the priors to penalize the model for inferring latent coordinates that result in distances that are comparable in scale to the observed covariates. This would embed a tendency for the model to use the observed covariates, rather than the latent space, to model the outcome. Consider the following example of a one-dimensional spatial model with an additional dyadic covariate.

$$pr(y_{ij} = 1) = \text{logit}^{-1}(\beta_0 + \beta_1 x_{ij} - |z_i - z_j|).$$

If we assume that the z are independent and distributed $N(0, \sigma^2)$, it is straightforward to derive the distribution of $|z_i - z_j|$.² Since it is a special case of a “normal difference distribution”, we know that $z_i - z_j \sim N(0, 2\sigma^2)$ (Devore and Berk 2012). The normality of $z_i - z_j$ implies that $|z_i - z_j|$ has a folded normal distribution (Leone, Nelson and Nottingham 1961). Following from this result, we know that the variance of $|z_i - z_j|$ is

$$\sigma^2(2 - 4/\pi).$$

Let α^2 be the empirical variance of x . Then the variance in the linear predictor is $\beta_1^2 \alpha^2$. Given an estimate $\hat{\beta}_1$, we can set the variance of the latent space prior to approximately

²The assumption that $z \sim N(0, \sigma^2)$ is consistent with z being drawn from the prior that is conventionally used in the LSM.

$\hat{\beta}_1^2 \alpha^2 / (2 - 4/\pi)$ in order to tune the prior to avoid replicating x via the distances between latent coordinates. This result is limited in that it requires the assumption of normally distributed latent coordinates and applies only to a single latent dimension. However, it may provide a reasonable approximation in the case of non-normal coordinates and/or multiple dimensions. Thirey and Hickman (2015) derive the distribution of Euclidean Distances between two k -dimensional points where each coordinate is drawn from a standard normal distribution.

4 Analysis

4.1 simulation design

To study how the LSM behaves in the situation in which we have some observed covariates as well as omitted network structure which can be represented using a Euclidean latent space. For simplicity we consider a unidimensional latent space. To generate an observed covariate which has a controllable collinearity with the latent network structure we follow a three-step process. First, we simulate k -dimensional positions for each node, drawing each from a normal distribution, and calculate dyadic distances d between each pair of positions. Second, given a target covariance matrix (Σ) among the covariates and distances, $\langle X, d \rangle$, we derive the conditional mean vector and covariance, assuming that X has a joint normal distribution given d (see (Eaton 1983, pp. 116–117) for the conditional normal derivation). Third, we simulate X as a multivariate normal random variable with the respective conditional means and covariance. Finally we standardize each X to have zero mean and unit variance. To generate the covariance matrix Σ which controls the dependence between the omitted network structure and the observed covariate X we utilize the C-vine method of Lewandowski, Kurowicka and Joe (2009).

We consider three exponential family distributions for the edges: Gaussian, Binomial, and Poisson. Additionally we control the number of nodes in the network $n = 25, 50, 100$, fix the dimension of the latent space k at 1 for simplicity, and consider 4 values of the collinearity parameter η , .1, 1, 100, and 1000000, where .1 corresponds to near perfect collinearity and 1000000 to independence. See Figure 1 for the distribution of the absolute value of the maximal off-diagonal correlation generated at each value of η .

Each step in the simulation involves the estimation of a GLM, and the LSM. We consider the LSM with several different priors on the coefficient for the observed covariate and the latent space. We set the prior variance of β_1 to be either 1 or 10 and alternatively use a diffuse normal prior on the latent space or scale it by $\hat{\beta}_1^2 \alpha^2 / (2 - 4/\pi)$, where α^2 is the empirical variance of X .

At each step in the simulation, the LSM is estimated using the canonical imple-

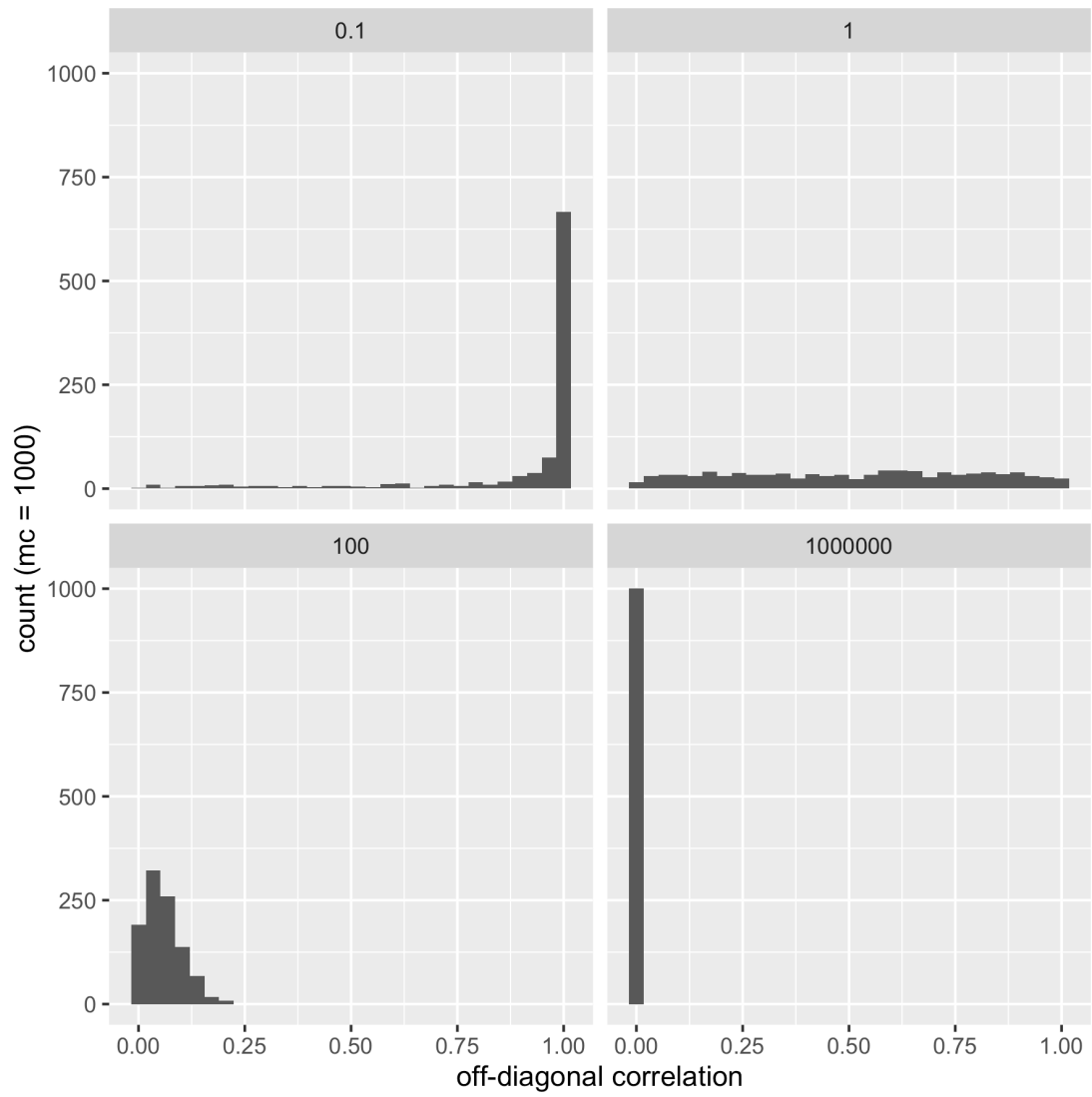


Figure 1: The distribution of the absolute value of the maximum off-diagonal correlation.

mentation in the `latentnet` package in R. An initial run of 10,000 burn in iterations, followed by 1,000,000 iterations of the sampler. Every 100th iteration is saved. Convergence in the log probability of the model is assessed using the Geweke diagnostic. If the convergence criteria is satisfied the simulation continues to the next set of arguments, otherwise the number of iterations is doubled. If the convergence criteria is still not satisfied, then the aforementioned step in the simulation is flagged for review. At each point in the simulation’s parameter space, we execute 500 Monte Carlo iterations.

One of the primary difficulties in executing the above simulation design is the computational cost of estimating the LSM. We utilize the `BatchJobs` and `BatchExperiments` R packages to construct and execute our computational experiments on a Torque cluster Bischl, Lang, Mersmann, Weihs et al. (2015).

We first compute the MLE estimated by iteratively reweighted least squares in R. For the LSM we use the posterior mode. In the cases where edges are Binomial distributed, we scale the estimates using the reciprocal of the bias of β_1 when X and d are independent: $\frac{\sqrt{3.28 + \beta_2^2 \text{Var}(d)}}{\sqrt{3.29}}$, where 3.29 is the variance of a standard logistic distribution (See the derivation of the bias under an independent but omitted covariate in Mood (2010)).

We consider networks of varying size (i.e. number of nodes) and collinearity structure between the omitted distance between nodes and the measured covariates. For each point in this space we evaluate the mean square prediction error of the model on new data drawn from the same distribution (i.e., the generalization error), the bias of the estimated coefficients for the measured covariate, the type 2 error rate (i.e., the probability that the null hypothesis that each regression parameter is 0 is incorrectly accepted), and the type 1 error rate, that is, the probability of rejecting the null hypothesis that the effect of the measured covariate is 0 when it is in fact 0. Inference regarding β_1 is computed using the maximum a posteriori estimate (i.e., the mode of the posterior), and credible intervals for β_1 are defined by using the region of highest posterior density which covers 95% of the marginal posterior distribution for each regression parameter.

To evaluate estimation error we compute the bias: the difference between the true coefficient β_1 and the estimate. Figures 2 and 3 shows the results. When a latent space is present (i.e., $\beta_2 = -1$, shown in Figure ??) the LSM outperforms the GLM in terms of bias across the board. However, when latent network structure is not present ($\beta_2 = 0$, shown in Figure ??) the LSM is biased upwards in the Binomial case and appears to have higher variance (perhaps due to efficiency losses) for Poisson and Gaussian edges (preliminary interpretation).

Figures 4 and 5 show the inferential error rates of the LSM with different priors for β_1 and the latent space compared to a GLM. Type 1 error rates for the LSM are in general lower than those of a GLM when an unmeasured covariate exists, however, when one does not exist the GLM outperforms the LSM. However, scaling the prior of the latent space to match that of the observed covariates appears to make the LSM’s error rate

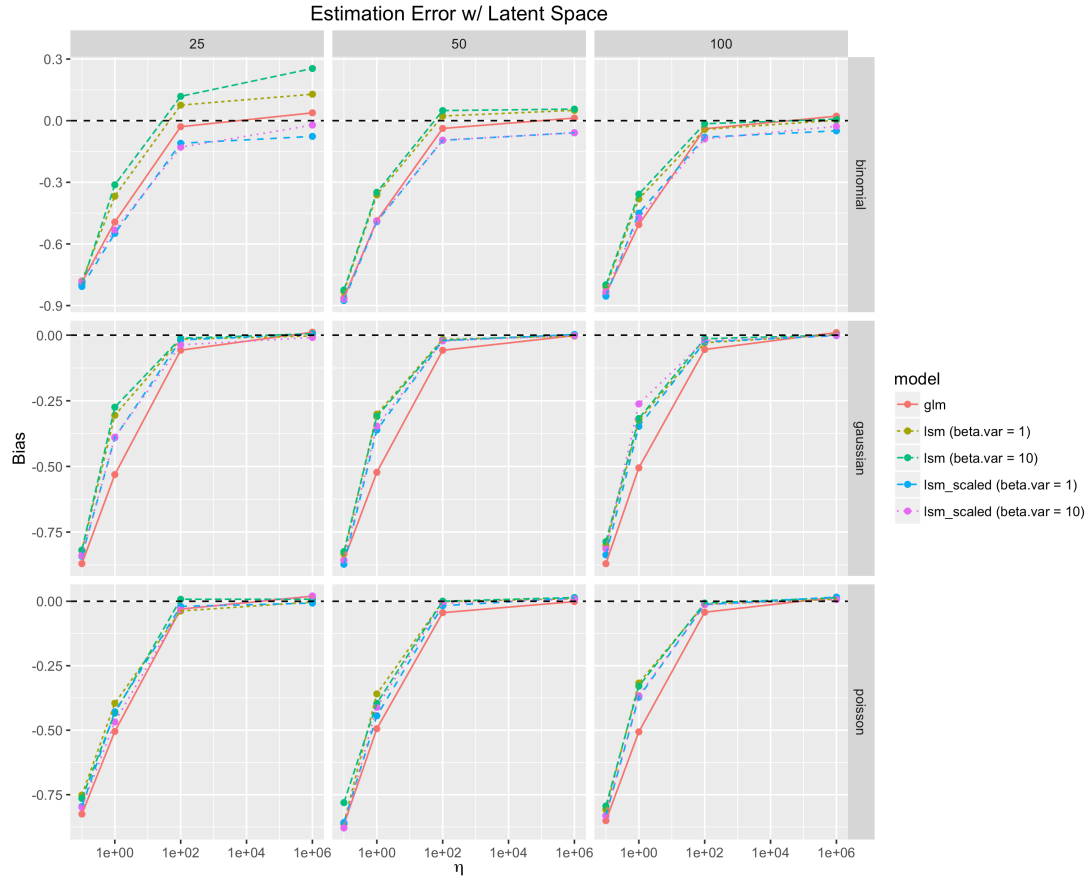


Figure 2: The bias of estimates of the effect of the observed covariate X when there is an omitted variable. The x -axis gives the value of the parameter η which controls the degree of dependence between X and omitted covariate. Lower values of η indicate higher levels of dependence between the observed and omitted covariate. The y -axis gives a Monte Carlo estimate of the bias. The number of nodes are indicated in the top panels, while the distributional family of the edges is shown on the right panel. Each panel represents 4 values of η with 500 (partially complete) Monte Carlo iterations executed at each point.

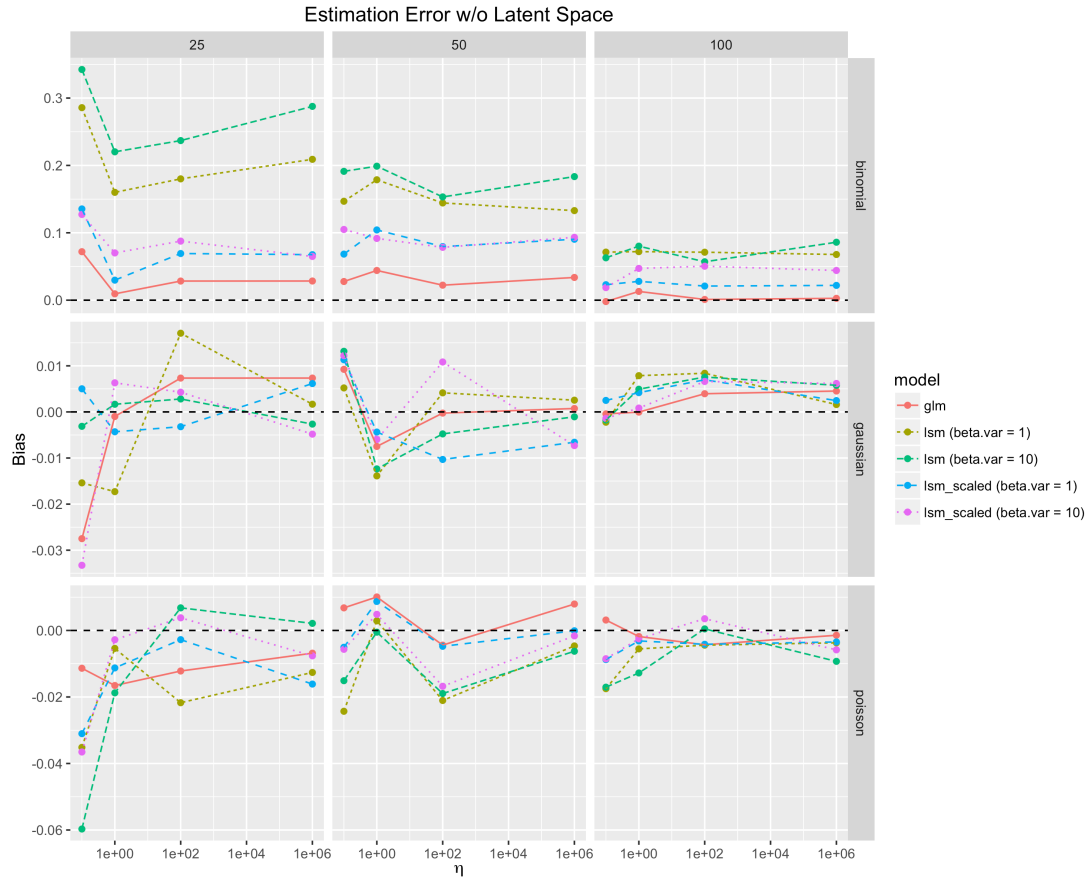


Figure 3: The bias of estimates of the effect of the observed covariate X when there is no omitted variable.

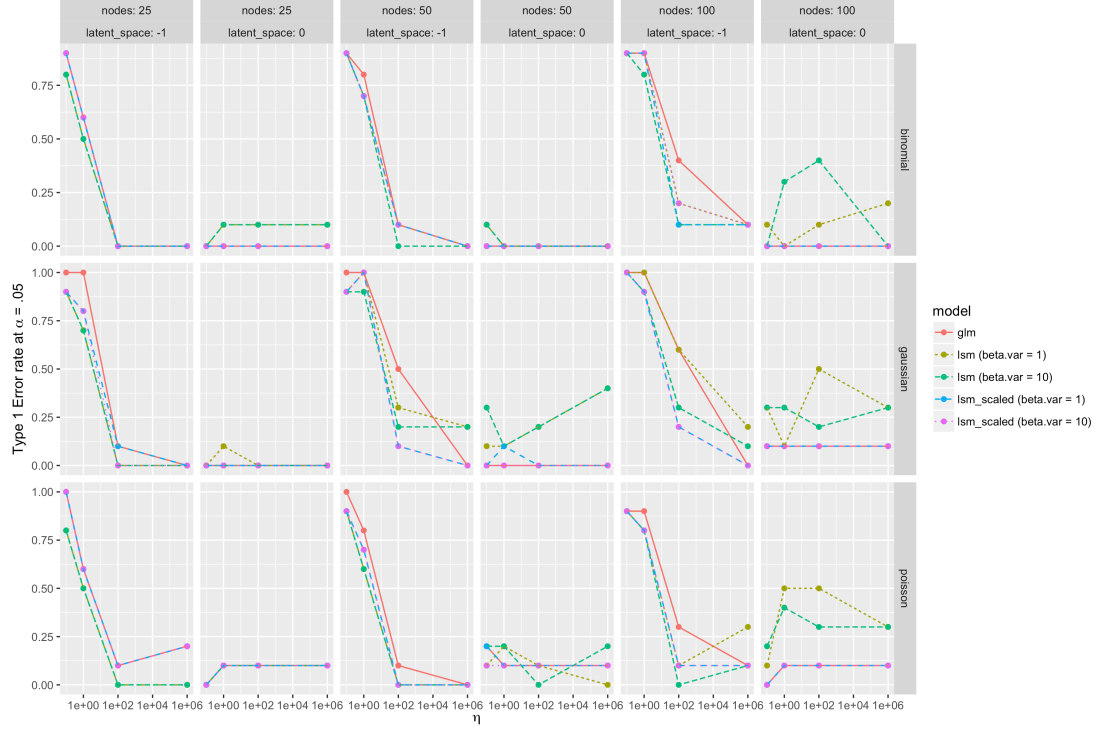


Figure 4: Monte Carlo estimates of the Type-1 error regarding the effect of the observed covariate X are shown on the y - axis. Here, $\beta_1 = 0$ and the error rate shown in each panel in that row gives 1 minus the probability of a 95% confidence region (for the LSM) or interval (for the GLM) including 0, giving the probability that a true null hypothesis of $\beta_1 = 0$ is falsely rejected.

comparable to that of the GLM. Type 2 error rates are very similar for the GLM and the LSM with all evaluated prior settings; to the point that they are indistinguishable.

Generalization error, in this case defined as the expected error on new edges generated using a fixed set of latent distances between the nodes and a node-level covariate is shown in figure 6. In nearly all cases the LSM substantially outperforms the GLM, especially when the omitted covariate is not highly collinear with the observed covariate.

5 Conclusion

Although the results are (very) preliminary (and have been plagued by my incompetence!), it seems increasingly clear that the primary reason for using the LSM or one of its many variants, rather than a GLM, should be interest in using the latent space to uncover patterns of interest or to predict edges in networks where there is likely to

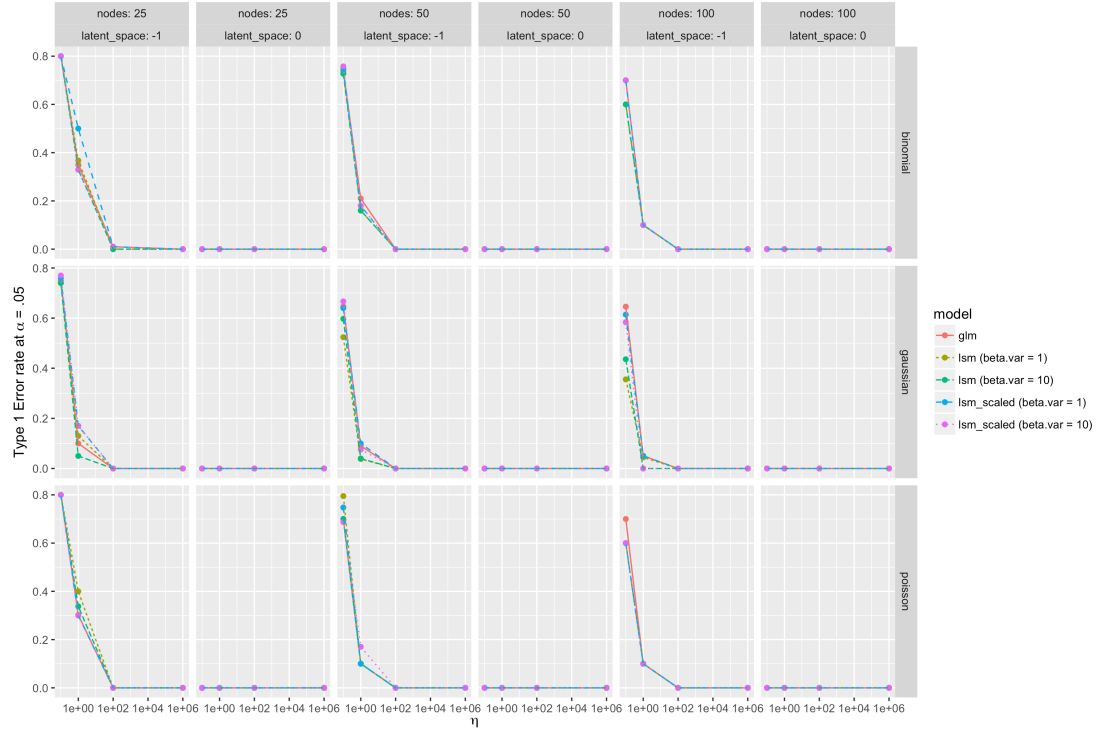


Figure 5: Monte Carlo estimates of the Type-2 error regarding the effect of the observed covariate X are shown on the y - $axis$. Here, $\beta_1 = 1$, and the error rate shown in each panel in each row gives the probability of the probability/confidence intervals covering 0, giving the probability of accepting a false null hypothesis.

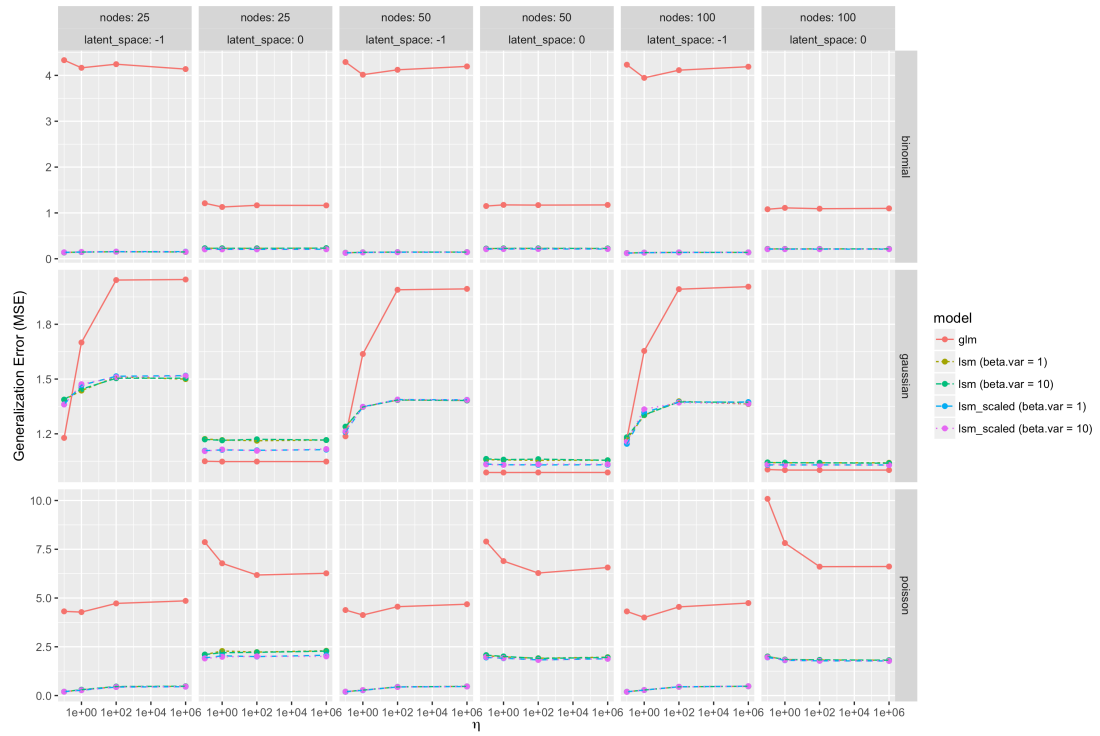


Figure 6: The x -axis shows a Monte Carlo estimate of the mean square error for edge values drawn from the generative model using a model trained on independent data.

be omitted structure that is not adequately modelled using measured covariates. However, inferential and estimation error are often smaller than that of a GLM when an omitted variable that corresponds to a form that can be modelled using a latent space is present, and can lead to an up to 25% decrease in bias relative to a GLM. Whether or not there is a reduction in estimation and/or inferential error depends on the strength of the omitted confounder, sample size (the LSM having many more parameters than an equivalent GLM), and relies on the omitted confounder being representable as a latent space. Lastly, there is some (preliminary) evidence to suggest that scaling the prior of the latent space to match that of the measured covariates can reduce Type-1 error.

References

- Bischl, Bernd, Michel Lang, Olaf Mersmann, Claus Weihs et al. 2015. “BatchJobs and BatchExperiments: Abstraction mechanisms for using R in batch environments.” *Journal of Statistical Software* 64(1):1–25.
- Cao, Xun and Michael D Ward. 2013. “Do democracies attract portfolio investment?” *International Interactions*, *forthcoming*.
- Chiu, Grace S and Anton H Westveld. 2011. “A unifying approach for food webs, phylogeny, social networks, and statistics.” *Proceedings of the National Academy of Sciences* 108(38):15881–15886.
- Dass, Mayukh and Piyush Kumar. 2011. “The impact of economic and social orientation on trust within teams.” *Journal of Business & Economics Research (JBER)* 9(2).
- Dass, Mayukh and Piyush Kumar. 2012. “Assessing category vulnerability across retail product assortments.” *International Journal of Retail & Distribution Management* 40(1):64–81.
- Davis, CE, JE Hyde, SI Bangdiwala and JJ Nelson. 1986. “An example of dependencies among variables in a conditional logistic regression.” *Modern statistical methods in chronic disease epidemiology* pp. 140–147.
- Devore, Jay L and Kenneth N Berk. 2012. Inferences based on two samples. In *Modern Mathematical Statistics with Applications*. Springer pp. 484–551.
- Eaton, Morris L. 1983. *Multivariate statistics: a vector space approach*. New York: Wiley.
- Fearnbach, Holly, John W Durban, David K Ellifrit, Janice M Waite, Craig O Matkin, Chris R Lunsford, Megan J Peterson, Jay Barlow and Paul R Wade. 2014. “Spatial and social connectivity of fish-eating Resident killer whales (*Orcinus orca*) in the northern North Pacific.” *Marine biology* 161(2):459–472.
- Fletcher Jr, Robert J, Andre Revell, Brian E Reichert, Wiley M Kitchens, Jeremy D Dixon and James D Austin. 2013. “Network modularity reveals critical scales for connectivity in ecology and evolution.” *Nature communications* 4.
- Fletcher, Robert J, Miguel A Acevedo, Brian E Reichert, Kyle E Pias and Wiley M Kitchens. 2011. “Social network models predict movement and connectivity in ecological landscapes.” *Proceedings of the National Academy of Sciences* 108(48):19282–19287.

- Handcock, Mark S, Adrian E Raftery and Jeremy M Tantrum. 2007. "Model-based clustering for social networks." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 170(2):301–354.
- Hoff, Peter. 2008. Modeling homophily and stochastic equivalence in symmetric relational data. In *Advances in Neural Information Processing Systems*. pp. 657–664.
- Hoff, Peter D. 2003. *Random effects models for network data*. na.
- Hoff, Peter D. 2009. "Multiplicative latent factor models for description and prediction of social networks." *Computational and mathematical organization theory* 15(4):261–272.
- Hoff, Peter D, Adrian E Raftery and Mark S Handcock. 2002. "Latent space approaches to social network analysis." *Journal of the american Statistical association* 97(460):1090–1098.
- Kirkland, Justin H. 2012. "Multimember Districts' Effect on Collaboration between US State Legislators." *Legislative Studies Quarterly* 37(3):329–353.
- Krivitsky, Pavel N., Mark S. Handcock, Adrian E. Raftery and Peter D. Hoff. 2009. "Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models." *Social Networks* 31(3):204 – 213.
URL: <http://www.sciencedirect.com/science/article/pii/S0378873309000173>
- Leone, FC, LS Nelson and RB Nottingham. 1961. "The folded normal distribution." *Technometrics* 3(4):543–550.
- Lewandowski, Daniel, Dorota Kurowicka and Harry Joe. 2009. "Generating random correlation matrices based on vines and extended onion method." *Journal of multivariate analysis* 100(9):1989–2001.
- Mackinnon, Murray J and Martin L Puterman. 1989. "Collinearity in generalized linear models." *Communications in statistics-theory and methods* 18(9):3463–3472.
- Mitchell, Toby J and John J Beauchamp. 1988. "Bayesian variable selection in linear regression." *Journal of the American Statistical Association* 83(404):1023–1032.
- Mood, Carina. 2010. "Logistic regression: Why we cannot do what we think we can do, and what we can do about it." *European Sociological Review* 26(1):67–82.
- Paul, Sudeshna, Nancy L Keating, Bruce E Landon and A James O'Malley. 2014. "Results from using a new dyadic-dependence model to analyze sociocentric physician networks." *Social Science & Medicine* 117:67–75.

- Shmueli, Galit. 2010. "To explain or to predict?" *Statistical science* pp. 289–310.
- Simpson, Sean L, F DuBois Bowman and Paul J Laurienti. 2013. "Analyzing complex functional brain networks: fusing statistics and network science to understand the brain." *Statistics surveys* 7:1.
- Thirey, Benjamin and Randal Hickman. 2015. "Distribution of Euclidean Distances Between Randomly Distributed Gaussian Points in n-Space." *arXiv preprint arXiv:1508.02238* .
- Vance, Eric A, Elizabeth A Archie and Cynthia J Moss. 2009. "Social networks in African elephants." *Computational and mathematical organization theory* 15(4):273–293.
- Ward, Michael D, John S Ahlquist and Arturas Rozenas. 2013. "Gravity's rainbow: a dynamic latent space model for the world trade network." *Network Science* 1(01):95–118.
- Ward, Michael D, Randolph M Siverson and Xun Cao. 2007. "Disputes, democracies, and dependencies: A reexamination of the Kantian peace." *American Journal of Political Science* 51(3):583–601.
- Weissfeld, Lisa A and Susan M Sereika. 1991. "A multicollinearity diagnostic for generalized linear models." *Communications in Statistics-Theory and Methods* 20(4):1183–1198.
- Westveld, Anton H and Peter D Hoff. 2011. "A mixed effects model for longitudinal relational and network data, with applications to international trade and conflict." *The Annals of Applied Statistics* pp. 843–872.
- Zhang, Zhaoyang, Honggang Wang, Chonggang Wang and Hua Fang. 2015. "Cluster-based Epidemic Control Through Smartphone-based Body Area Networks." *Parallel and Distributed Systems, IEEE Transactions on* 26(3):681–690.