G Model
SON-1034; No. of Pages 10
ARTICLE IN PRESS
Social Networks xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Social Networks
journal homepage: www.elsevier.com/locate/socnet

Heterogeneous susceptibilities in social influence models

Daniel K. Sewell

Department of Biostatistics, University of Iowa, Iowa City, IA 52242, United States

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Dependent data
Durbin model
Egocentric network
Hierarchical model
Laplace approximation
Network autocorrelation model

ABSTRACT

Network autocorrelation models are widely used to evaluate the impact of social influence on some variable of interest. This is a large class of models that parsimoniously accounts for how one's neighbors influence one's own behaviors or opinions by incorporating the network adjacency matrix into the joint distribution of the data. These models assume homogeneous susceptibility to social influence, however, which may be a strong assumption in many contexts. This paper proposes a hierarchical model that allows the influence parameter to be a function of individual attributes and/or of local network topological features. We derive an approximation of the posterior distribution in a general framework that is applicable to the Durbin, network effects, network disturbances, or network moving average autocorrelation models. The proposed approach can also be applied to investigating determinants of social influence in the context of egocentric network data. We apply our method to a data set collected via mobile phones in which we determine the effect of social influence on physical activity levels, as well as classroom data in which we investigate peer influence on student defiance. With this last data set, we also investigate the performance of the proposed egocentric network model.

1. Introduction

Social influence can help explain behaviors, opinions and beliefs, and as such is of importance in sociology, marketing, public health, political science, etc. Newcomb (1951) describes social influence thus:

> Any observable behavior [e.g., a displayed position on an issue] is not only a response (on the part of a subject) which is to be treated as a dependent variable; it is also a stimulus to be perceived by others with whom the subject interacts, and thus to be treated as an independent variable.

A prodigious amount of research has been done evaluating the effects of social influence on various attributes. Examples include social influence on binge eating (Crandall, 1988), smoking and drinking behaviors of youth (Simons-Morton et al., 2001), investment decisions (Hoffmann and Broekhuizen, 2009), emotions (Hareli and Rafaeli, 2008), and transitioning from noninjecting heroine user to an injecting user (Neaigus et al., 2006).

The statistical models of choice for accounting for and evaluating the impact of social influence has long been the class of network autocorrelation models. These have been deemed the 'workhorse for modeling network influences on individual behavior' (Fujimoto et al., 2011). These models have their roots in spatial statistics, with important early works by Ord (1975) and Doreian (1980). These same models were quickly and successfully used for network data, and are still widely used and studied.

Network autocorrelation models make the strong assumption of homogeneous susceptibilities to social influence among network members; that is, each actor in the network is assumed to be equally susceptible to peer influence. This is contrary to much of the research being done by substantive scientists. Friedkin and Johnsen (1999) devised a theory of social influence that incorporates heterogeneity in susceptibility to influences through the network. Empirical studies have also shown that susceptibility can in fact vary based on subject attributes. For example, studies have shown that birth order (Staples and Walters, 1961), age (Krosnick and Alwin, 1989; Steinberg and Monahan, 2007), and gender (Eagly, 1978) all affect an individual's resistance to social influence; Fennis and Aarts (2012) show that reducing personal control leads to a higher susceptibility to social influence; Urberg et al. (2003) show that relationship variables can increase conformity to one's peers with respect to substance-use in adolescents; countless papers have been published using the Consumer Susceptibility to Interpersonal Influence measure (Bearden et al., 1989), determining a subject's susceptibility to social influence.

Using network autocorrelation models as a starting point, we relax the assumption of uniform susceptibility to social influence. This is done by using a non-linear hierarchical model in which an

E-mail address: daniel-sewell@uiowa.edu

http://dx.doi.org/10.1016/j.socnet.2017.06.004

individual's susceptibility is a function of that individual's characteristics. These characteristics can take on a variety of forms; the most obvious of these is an individual's attributes such as gender, but one may also use local network topological features, such as centrality measures, and in so doing determine how the network topology itself can affect how an actor may be susceptible to social influence. The proposed model can follow from the major network autocorrelation models, namely the Durbin model, the network effects model, the network disturbances model, and the network moving average model. Just as in homogeneous network autocorrelation models, the posterior distribution for our model is available in closed form but is not a well-known distribution. However, by using the Laplace approximation, we show how to very quickly obtain approximate samples from the posterior.

Section 2 describes our proposed methods and our approximation to the posterior distribution. In Section 3, we extend these ideas to egocentric network data. Section 4 describes a simulation study. Section 5 describes an analysis of data collected via mobile phones measuring call logs among network members and health activity. Section 6 describes an analysis of educational data collected via surveys measuring student attitudes and behaviors. We end with a brief conclusion in Section 7.

2. Methods

In this section, we first provide a brief review of the four most common types of network autocorrelation models. We then describe the proposed model and Bayesian estimation procedure for the simplest of these, which assumes a simple diagonal covariance matrix for the response vector. Next, we describe more sophisticated and realistic social influence models and how to adapt the estimation scheme accordingly. We end the section by relating our work to a common operationalization of the adjacency matrix, namely row normalization.

Before beginning with the review, we shall provide some notation. We will let $\boldsymbol{y}$ be the $n \times 1$ response vector of interest, where n is the total number of actors in the network. Let X_1 and X_2 be design matrices corresponding to the independent variables, having dimension $n \times p_1$ and $n \times p_2$ respectively. Let A be the $n \times n$ adjacency matrix such that $A_{ij} = 1$ if there is an edge from actor i to actor j and zero otherwise; $A_{ii} = 0$ for all i. In some cases A may represent a weighted network where the non-diagonal entries of A may take values in some subset of $\Re$; these values typically represent the strength of the edges in some meaningful way, e.g., the count of interactions between i and j. Let W_x and W_ϵ be design matrices with dimension $n \times q_1$ and $n \times q_2$ respectively constructed from independent variables, functions of local network topologies, or some combination of the two. Note that X_1, X_2, W_x, and W_ϵ may be identical, share some covariates, or be constructed from entirely different covariates. We will use $\mathbf{diag}(A)$ to represent the column vector constructed from the diagonal entries of some square matrix A, and $\text{Diag}(\mathbf{a})$ to represent the diagonal matrix constructed by setting the diagonal elements to be the elements of the vector $\mathbf{a}$. Sometimes it will be necessary to refer to a row, column, or element of a matrix; for any matrix M this will be denoted by $M_{(i,\cdot)}$, $M_{(\cdot,j)}$, or $M_{(i,j)}$ respectively.

2.1. Review of network autocorrelation models

The class of network autocorrelation models provide a solid statistical framework with which to investigate the effects of social influence, or, should social influence be considered a nuisance parameter, appropriately account for the complex dependencies in the data due to the network effect. This class is typically associated with four statistical models. The simplest of these is the Durbin model. This model assumes that the observations are independent given the network and the covariates, but that an individual's mean is affected by the covariates of his/her neighbors. Specifically, the Durbin model is given by

$$\boldsymbol{y} = X_1\boldsymbol{\beta}_1 + \rho_x A X_2 \boldsymbol{\beta}_2 + \boldsymbol{\epsilon}, \tag{1}$$

where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are parameter vectors of unknown coefficients of size p_1 and p_2 respectively, ρ_x is the parameter that captures the (uniform) social influence effect, and $\boldsymbol{\epsilon}$ is a vector of independent mean zero normal random variables with variance σ^2. Note that ρ_x is constrained to equal 1 for model identifiability. The assumption of independence is most often unreasonable in the context of network data. To address this, the Durbin model can be augmented to allow for correlated errors. Three ways to do this are the effects, disturbances, and moving average models (see, e.g., Doreian, 1980; Hepple, 1995).

The network effects model is given by

$$\boldsymbol{y} = \rho_\epsilon A\boldsymbol{y} + X_1\boldsymbol{\beta}_1 + \rho_x A X_2 \boldsymbol{\beta}_2 + \boldsymbol{\epsilon}. \tag{2}$$

In this model, in addition to the effect of neighbors' covariates, an individual's mean response is a function of his/her neighbors' responses. The network disturbances model is given by

$$\begin{aligned} \boldsymbol{y} &= X_1\boldsymbol{\beta}_1 + \rho_x A X_2 \boldsymbol{\beta}_2 + \boldsymbol{\nu}, \\ \boldsymbol{\nu} &= \rho_\epsilon A \boldsymbol{\nu} + \boldsymbol{\epsilon}. \end{aligned} \tag{3}$$

Hence the network disturbances model is the Durbin model with the network effects model (sans covariates) on the errors. This model can be interpreted to mean that an individual's deviation from his/her mean is a function of his/her neighbors' deviations from their mean. Similar in spirit is the network moving average model, given by

$$\boldsymbol{y} = X_1\boldsymbol{\beta}_1 + \rho_x A X_2 \boldsymbol{\beta}_2 + \boldsymbol{\epsilon} + \rho_\epsilon A \boldsymbol{\epsilon}. \tag{4}$$

The errors are additive based on the network structure, and hence an individual's response depends on the random fluctuations of his/her neighbors that cannot be explained by the neighbors' covariates.

2.2. Durbin model

Here we relax the assumption of homogeneous susceptibility to social influence, beginning with the Durbin model. If an individual's susceptibility is unique to them, then we may rewrite (1) as

$$\boldsymbol{y} = X_1\boldsymbol{\beta}_1 + R_x A X_2 \boldsymbol{\beta}_2 + \boldsymbol{\epsilon},$$

where R_x is some diagonal matrix. If we further hypothesize that an individual's susceptibility is determined by some set of covariates or local network functions contained in the design matrix W_x, we may let

$$\mathbf{diag}(R_x) = \tilde{W}_x \tilde{\boldsymbol{\gamma}}_x, \tag{5}$$

$$\tilde{W}_x = (\mathbf{1}, W_x), \tag{6}$$

$$\tilde{\boldsymbol{\gamma}}_x = (\mathbf{1}, \boldsymbol{\gamma}_x')', \tag{7}$$

where $\mathbf{1}$ is the vector of 1's. The constraints given in (6) and (7) ensure that the model is identifiable, under the (obvious) assumptions that no columns of W_x are proportional to $\mathbf{1}$ and all design matrices are of full rank.

We assume the priors on the parameters are of the following form:

$$\boldsymbol{\beta} \sim N(\mathbf{0}, g_1\sigma^2 I_{p_1+p_2}), \tag{8}$$

$$\boldsymbol{\gamma}_x \sim N(\mathbf{0}, g_2\sigma^2 I_{q_1}), \tag{9}$$

Please cite this article in press as: Sewell, D.K., Heterogeneous susceptibilities in social influence models. Soc. Netw. (2017), http://dx.doi.org/10.1016/j.socnet.2017.06.004