Computational Statistics and Data Analysis 54 (2010) 1657–1671

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Assessing local model adequacy in Bayesian hierarchical models using the partitioned deviance information criterion

David C. Wheeler [a,*], DeMarc A. Hickson [b], Lance A. Waller [c]

[a] National Cancer Institute, 6120 Executive Boulevard, Bethesda, MD 20892, United States
[b] University of Mississippi Medical Center, School of Medicine, Department of Medicine, 2500 North State Street, Jackson, MS 39213, United States
[c] Emory University, Rollins School of Public Health, Department of Biostatistics, 1518 Clifton Road, Atlanta, GA 30322, United States

ARTICLE INFO

Article history:
Received 11 August 2008
Received in revised form 20 January 2010
Accepted 22 January 2010
Available online 1 February 2010

Keywords:
Bayesian statistics
DIC
Spatial statistics
Hierarchical models
Linear models
HIV
Rwanda

ABSTRACT

Many diagnostic tools and goodness-of-fit measures, such as the Akaike information criterion (AIC) and the Bayesian deviance information criterion (DIC), are available to evaluate the overall adequacy of linear regression models. In addition, visually assessing adequacy in models has become an essential part of any regression analysis. In this paper, we focus on a spatial consideration of the local DIC measure for model selection and goodness-of-fit evaluation. We use a partitioning of the DIC into the local DIC, leverage, and deviance residuals to assess the local model fit and influence for both individual observations and groups of observations in a Bayesian framework. We use visualization of the local DIC and differences in local DIC between models to assist in model selection and to visualize the global and local impacts of adding covariates or model parameters. We demonstrate the utility of the local DIC in assessing model adequacy using HIV prevalence data from pregnant women in the Butare province of Rwanda during the period 1989–1993 using a range of linear model specifications, from global effects only to spatially varying coefficient models, and a set of covariates related to sexual behavior. Results of applying the diagnostic visualization approach include more refined model selection and greater understanding of the models as applied to the data.

Published by Elsevier B.V.

1. Introduction

Many diagnostic tools are available to evaluate the adequacy of a linear model. Model residuals are used to assess the model fit, while diagnostics such as leverage values, Cook's distances, DFFITS, and DFBETAS are used to identify outlying and influential observations. Residuals provide a well-known tool for identifying outlying data points and summarizing the contribution of each individual observation to the overall fit of a model, thus providing valuable elements for constructing model diagnostics. To aid in the evaluation of a model, diagnostic values are frequently presented in scatter plots with fitted values or covariates to identify observations that may be suspect according to one or more model characteristic (Neter et al., 1996). In fact, visually assessing the lack of fit in models has become an essential part of any regression analysis. This is evident in many diagnostic works in statistics, including that of Cook and Weisberg (1999), who use model checking plots to evaluate the appropriateness of the linear model. The model checking plot is a scatter plot of the fitted outcome and a function of the predictors, along with the ordinary least squares fit and a Lowess (locally weighted scatterplot smoothing) fit. Cook and Weisberg (1994) also use added-variable plots, scatter plots for visually assessing whether a variable has

* Corresponding author. Tel.: +1 301 435 4702; fax: +1 301 480 2669.
E-mail address: wheelerdc@mail.nih.gov (D.C. Wheeler).

0167-9473/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.csda.2010.01.025

explanatory power when added to the regression model of the outcome on another variable. These added-variable plots, along with ARES (Adding REgressors Smoothly) plots (Cook and Weisberg, 1994), are useful diagnostic plots for model selection. For visually identifying influential observations, Cook's distances are plotted against predictor values or are used to highlight certain observations in a plot of residuals versus predictors in generalized additive models (Hastie and Tibshirani, 1990). In spatial analyses, residuals themselves are also often mapped over the study unit to inspect for significant spatial autocorrelation of errors, a violation of the independence assumption of residuals in a linear model.

In addition to model diagnostic tools assessing the impact of individual observations, methods of assessing the overall goodness of fit and model complexity also have been developed for linear models, such as the Akaike information criterion (Akaike, 1973). The AIC is defined as AIC $= D(\hat{\theta}) + 2k$, the combination of the deviance evaluated at the maximum likelihood estimate of the parameters θ and a penalty defined as twice the number of model parameters. The deviance $D(\hat{\theta})$ is a general measure of fit defined as $D(\hat{\theta}) = -2 \log p(y|\hat{\theta})$, with $\log p(y|\hat{\theta})$ denoted as the maximized log likelihood. The AIC fits into the broad literature of classical covariate selection and model choice (Burnham and Anderson, 2002). There are also a variety of statistical assessments of overall model fit in the Bayesian paradigm, including Bayes factors (Kass and Raftery, 1995), the Bayesian information criterion (BIC: Schwarz, 1978) and the deviance information criterion (DIC: Spiegelhalter et al., 2002), among others (Gelman and Pardoe, 2006). As a generalization of the AIC, the DIC is appropriate for model comparison in complex hierarchical models where the number of parameters is unknown, such as the models used in spatial analysis, with disease mapping examples found in Zhu and Carlin (2000) and Best et al. (1999). As with the AIC, the DIC is a measure of model fit or adequacy with a penalty for model complexity. An advantage of the DIC is that one can easily calculate it from the Markov chain Monte Carlo (MCMC) simulation samples generated when drawing samples from the posterior distribution of a parameter in a Bayesian model. Another key advantage of the DIC is that it can be partitioned into individual contributions from observations in the data, as few diagnostic tools currently exist in the statistical literature to assess the local importance of additional model covariates within groups or subsets of data. The partitioning of the DIC into individual data contributions is outlined in Spiegelhalter et al. (2002). While the work of Spiegelhalter et al. (2002) provides the components necessary for an approach to local diagnosis of Bayesian model fit, an approach for visually diagnosing local spatial model fit has not been previously explored.

A local partitioning of the DIC is especially relevant in spatial model applications, such as in disease mapping, where relatively strong priors concerning spatial correlation among study units are often used. In the applied spatial statistical literature, there has been a recent emphasis on methodology that models local covariate effects, often spatially correlated, instead of the more traditional models that represent relationships with fixed effects across a study area. In reality, the association between a covariate and the occurrence of an outcome may vary between geographic and demographic subsets of individuals. Examples of reported differences in covariate and health outcome associations within a study population include: (1) prevalence of Hepatitis C virus among drug users across two geographic regions in Belgium (Mathei et al., 2004), (2) engagement in high-risk sexual behavior among men who have sex with men within different subgroups and geographic locations throughout the United States (McFarland et al., 2001; Leone et al., 2004; Guenther-Grey et al., 2005), and (3) prevalence of human immunodeficiency virus (HIV) among injecting drug users in urban and rural Scotland (Haw and Higgins, 1998). Such studies suggest that the impact of a determinant within a subset of individuals or a subset of the study area may drive an overall significant association and falsely suggest an association for the entire population. Conversely, parameter estimates based on the entire population may mask an influential impact limited to a subset of the population. Some methodology in spatial statistics recognizing the potential for these situations and allowing for regression relationships that vary over space are Bayesian spatially varying coefficient (SVC) models (Gelfand et al., 2003; Banerjee et al., 2004) and geographically weighted regression (Fotheringham et al., 2002). In practice, spatially varying coefficient models can be computationally demanding to fit, and local diagnostic tools that justify the additional computational effort in terms of improved local fit and more accurate representation of relationships are needed. In addition, local diagnostic methods to identify situations of differential fit and influence among spatial subgroups are currently not well developed. Related to this is subset analysis, which examines the effects of two or more treatments within each of several subsets, or subgroups, of data along with an overall assessment (Shafer and Olkin, 1983). An issue in subgroup analysis is that a large number of subgroups may be identified within a typical data set, which raises concern about multiplicity effects (Dixon and Simon, 1991). Partitioning of the DIC may be helpful in identifying areas or data subgroups where a specified spatial model is ill fitting or not particularly appropriate, i.e. where data are not in agreement with the prior.

In this paper, we focus on a spatial consideration of local DIC statistics for model selection and goodness-of-fit evaluation. We expand on the DIC partitioning approach of Spiegelhalter et al. (2002) to explore its applicability to visually assessing and quantifying the local model fit with individual and groups of spatial data. We use a partitioning of the DIC to assess the local model fit and data influence in a Bayesian framework for both individual observations and groups of observations, with groups corresponding to predefined spatial units. The interest in a partitioned DIC is to identify whether some models fit differently in certain areas and highlight any local, rather than global, impacts of covariate effects. In our approach to local model diagnosis, we introduce mapping of the partitioned, or local, DIC to explore spatial patterns of model fit over different spatial areas. We also map differences in local DIC values between models to examine the impact of adding additional covariates or model parameters. In addition, we plot the local DIC components of deviance residuals and leverage values and link plots of DIC components to maps of local DIC values. The novelty of this work involves building diagnostic tools out of available components in a typical Bayesian analysis to strengthen the data analysis, an area of rich results from linear models