Computational Statistics and Data Analysis 55 (2011) 1041–1052

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Regularized reduced rank growth curve models

Yoshio Takane *, Kwanghee Jung, Heungsun Hwang

McGill University, Department of Psychology, 1205 Dr. Penfield Avenue, Montreal, QC, H3A 1B1, Canada

ARTICLE INFO

Article history:
Received 8 November 2009
Received in revised form 11 August 2010
Accepted 17 August 2010
Available online 8 September 2010

Keywords:
The growth curve model (or GMANOVA)
Reduced rank approximation
Ridge-type regularization
A mixture of GMANOVA and MANOVA
Generalized singular value decomposition (GSVD)
Permutation tests
K-fold cross validation
The bootstrap method

ABSTRACT

The growth curve model (GCM), also known as GMANOVA, is a useful technique for investigating patterns of change in repeated measurement data over time and examining the effects of predictor variables on temporal trajectories. The reduced rank feature had been introduced previously to GCM for capturing redundant information in the criterion variables in a parsimonious way. In this paper, a ridge type of regularization was incorporated to obtain better estimates of parameters. Separate ridge parameters were allowed in column and row regressions, and the generalized singular value decomposition (GSVD) was applied for rank reduction. It was shown that the regularized estimates of parameters could be obtained in closed form for fixed values of ridge parameters. Permutation tests were used to identify the best dimensionality in the solution, and the K-fold cross validation method was used to choose optimal values of the ridge parameters. A bootstrap method was used to assess the reliability of parameter estimates. The proposed model was further extended to a mixture of GMANOVA and MANOVA. Illustrative examples were given to demonstrate the usefulness of the proposed method.

1. Introduction

The growth curve model (GCM; Potthof and Roy, 1964), also known as generalized MANOVA (or GMANOVA for short), is a useful technique for investigating patterns of change in repeated measurements of a response variable or variables over time and examining the effects of predictor variables on temporal trajectories. This type of model is often used in the analysis of longitudinal or repeated measurement data, often arising in psycho-physiological, biological and medical research. Recently, the reduced rank feature was introduced to GCM (Reinsel and Velu, 1998, 2003) to capture redundant information in the criterion variables in a parsimonious way. This additional feature allows the extraction of components of predictor variables that are most predictive of criterion variables. A series of components called redundancy components are mutually orthogonal and successively account for the maximum variance in the criterion variables.

In experimental studies conducted in biomedicine and psychology, we frequently encounter data with small sample sizes, which tend to produce estimates of parameters with large standard errors. The small sample size problem casts serious doubts about the adequacy of conventional estimation methods, such as the maximum likelihood estimation method, that largely rely on an asymptotic rationale. To remedy this situation, we incorporate a ridge type of regularization in estimating parameters in the reduced rank GCM. This method shrinks estimates of parameters toward zero, thereby reducing the variance of the estimates a great deal, while introducing a small bias. The net result is that estimates of parameters closer to true population values may be obtained. A ridge type of regularization method is particularly attractive when the sample size is small and/or predictor variables are nearly collinear (Hoerl and Kennard, 1970). This has been demonstrated recently in a variety of contexts in multivariate analysis (Hwang, 2009; Takane and Hwang, 2007; Takane et al., 2008; Takane and

* Corresponding author. Tel.: +1 514 398 6125; fax: +1 514 398 4896.
E-mail addresses: takane@psych.mcgill.ca, takane@takane2.psych.mcgill.ca (Y. Takane), kwanghee.jung@mail.mcgill.ca (K. Jung), heungsun.hwang@mcgill.ca (H. Hwang).

0167-9473/$ – see front matter
doi:10.1016/j.csda.2010.08.012

Jung, 2008, 2009). In this paper, we extend the basic methodology of ridge regularization to the reduced rank GCM and illustrate its use. We also consider an analogous extension of a mixture of the GMANOVA and MANOVA models (Chinchilli and Elswick, 1985) with the GMANOVA part subject to similar rank reduction, and the MANOVA part capturing the effects of extraneous variables.

This paper is organized as follows. We first present the model and the parameter estimation procedure for the regularized reduced rank GCM (Section 2.1). We then extend the model and the estimation procedure to a mixture of the GMANOVA and MANOVA models (Section 2.2). This is followed by expositions of permutation tests for selecting the best dimensionality in the solution, the K-fold cross validation method for choosing optimal values of ridge parameters, and the bootstrap method for assessing the reliability of parameter estimates (Section 2.3). Illustrative examples are given to demonstrate the usefulness of the proposed method in simulated and real data analysis situations (Section 3). The final section concludes the paper (Section 4).

2. The methods

2.1. The reduced rank GCM and the regularized parameter estimation

Let $\mathbf{Y}$ denote an n by p matrix of criterion variables. In the GCM setting, this matrix typically consists of multiple measurements of a response variable at p time points from a group of n subjects or cases, although in more general settings, it could be any multivariate data matrix. We assume that there is some additional information about the subjects and/or about the variables in $\mathbf{Y}$ that may be used to predict parts of $\mathbf{Y}$. Let $\mathbf{X}$ denote an n by q ($q \le n$) matrix of predictor variables for subjects such as their group memberships (e.g., treatment groups) and other demographic information. Let $\mathbf{H}$ denote a p by d ($d \le p$) matrix of predictor variables for time points (or variables) in $\mathbf{Y}$ that capture the relationships among the columns of $\mathbf{Y}$ such as the coefficients of orthogonal polynomials over time. (The matrix $\mathbf{X}$ is often called a between-subjects design matrix, and $\mathbf{H}$ a within-subjects design matrix.) Then, the GCM may be written as

$$\mathbf{Y} = \mathbf{XBH}' + \mathbf{E}, \tag{1}$$

where $\mathbf{B}$ is a q by d matrix of regression coefficients, and $\mathbf{E}$ is an n by p matrix of disturbance terms. In the reduced rank GCM, we assume that there is some redundancy in $\mathbf{B}$, so that

$$\text{rank}(\mathbf{B}) = r \le \min(q, d) \tag{2}$$

(Reinsel and Velu, 1998, 2003). A model of the above form has existed outside the realm of GCM, e.g., 2-way CANDELINC (CANonical DEcomposition under LINnear Constraints Carroll et al., 1980), and as a special case of CPCA (Constrained Principal Component Analysis; Takane and Shibayama, 1991; Takane and Hunter, 2001). Note that, if there is no obvious $\mathbf{H}$ available, we set $\mathbf{H} = \mathbf{I}$, and the model reduces to a simple MANOVA or redundancy analysis model (Van den Wollenberg, 1977; van der Leeden, 1990).

Parameters in the reduced rank GCM are usually estimated by the maximum likelihood (ML) method (Reinsel and Velu, 1998, 2003) or by the least squares (LS) method (Carroll et al., 1980; Takane and Shibayama, 1991; Takane and Hunter, 2001). We use the latter with the provision of its extension to regularized estimation in mind. The structure of derivations for the regularized estimation is remarkably similar to that for the non-regularized case. In the ordinary LS estimation, we minimize

$$\phi(\mathbf{B}) = \text{SS}(\mathbf{Y} - \mathbf{XBH}') \tag{3}$$

with respect to $\mathbf{B}$ subject to the rank restriction (2). To achieve this goal, we first rewrite $\phi(\mathbf{B})$ as (Takane and Shibayama, 1991; ten Berge, 1993):

$$\begin{aligned}\phi(\mathbf{B}) &= \text{SS}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\mathbf{H}') + \text{SS}(\hat{\mathbf{B}} - \mathbf{B})_{X'X,\,H'H} \\ &= \text{SS}(\mathbf{Y}) - \text{SS}(\mathbf{Y})_{P_X,\,P_H} + \text{SS}(\hat{\mathbf{B}} - \mathbf{B})_{X'X,\,H'H},\end{aligned} \tag{4}$$

where $\text{SS}(\mathbf{A})_{M,N} = \text{tr}(\mathbf{A}'\mathbf{MAN})$, $\mathbf{P}_X = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$ and $\mathbf{P}_H = \mathbf{H}(\mathbf{H}'\mathbf{H})^{-}\mathbf{H}'$ are orthogonal projectors onto the column spaces of $\mathbf{X}$ and $\mathbf{H}$, respectively, and

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{YH}(\mathbf{H}'\mathbf{H})^{-}, \tag{5}$$

is a rank free LS estimate of $\mathbf{B}$. Here "$^-$" indicates a generalized inverse (g-inverse). Note that while $\hat{\mathbf{B}}$ in (5) is not unique if $\mathbf{X}$ or $\mathbf{H}$ is singular, the decomposition (4) is unique. To obtain a unique estimate of $\hat{\mathbf{B}}$, we can use the Moore–Penrose inverse for $(\mathbf{X}'\mathbf{X})^-$ and $(\mathbf{H}'\mathbf{H})^-$. Since the first and the second term on the right-hand side of (4) are unrelated to $\mathbf{B}$, the reduced rank estimate of $\mathbf{B}$ can be obtained by minimizing the third term. This can be done via the generalized singular value decomposition (GSVD) of $\hat{\mathbf{B}}$ with metric matrices $\mathbf{X}'\mathbf{X}$ and $\mathbf{H}'\mathbf{H}$. This GSVD problem is written as

$$\text{GSVD}(\hat{\mathbf{B}})_{X'X,\,H'H}. \tag{6}$$