Journal of Multivariate Analysis 145 (2016) 87–100

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Modeling the Cholesky factors of covariance matrices of multivariate longitudinal data

Priya Kohli [a,*], Tanya P. Garcia [b], Mohsen Pourahmadi [c]

[a] *Department of Mathematics, Connecticut College, 270 Mohegan Avenue, New London, CT 06320, United States*

[b] *Department of Epidemiology and Biostatistics, Texas A&M Health Science Center, 1266 TAMU, College Station, TX 77843-1266, United States*

[c] *Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143, United States*

ARTICLE INFO

Article history:
Received 4 June 2015
Available online 14 December 2015

AMS subject classifications:
62G
62H
62-09

Keywords:
Cholesky decomposition
Linear covariance model
Majorization–minimization
Principal component analysis
Quadratic programming

ABSTRACT

Modeling the covariance matrix of multivariate longitudinal data is more challenging as compared to its univariate counterpart due to the presence of correlations among multiple responses. The modified Cholesky *block* decomposition reduces the task of covariance modeling into parsimonious modeling of its two matrix factors: the regression coefficient matrices and the innovation covariance matrices. These parameters are statistically interpretable, however ensuring positive-definiteness of several (innovation) covariance matrices presents itself as a new challenge. We address this problem using a subclass of Anderson's (1973) *linear covariance models* and model several covariance matrices using linear combinations of known positive-definite basis matrices with unknown non-negative scalar coefficients. A novelty of this approach is that positive-definiteness is guaranteed by construction; it removes a drawback of Anderson's model and hence makes linear covariance models more realistic and viable in practice. Maximum likelihood estimates are computed using a simple iterative majorization–minimization algorithm. The estimators are shown to be asymptotically normal and consistent. Simulation and a data example illustrate the applicability of the proposed method in providing good models for the covariance structure of a multivariate longitudinal data.

1. Introduction

Multivariate longitudinal studies involve two or more outcomes of interest measured repeatedly at two or more time points. These studies are common in clinical trials and biological research. The covariance matrix for such data plays a prominent role in analysis as it serves as a measure of temporal and cross-sectional dependence. Compared to univariate longitudinal data, developing a statistically interpretable and computationally efficient covariance model is more challenging in the multivariate case. This is because of the *positive-definiteness* constraint on the covariance matrix and *high-dimensionality* where now the number of parameters grows quadratically with the number of outcomes and time points. Graphical tools to visualize dependence patterns may also involve a multitude of graphs. Therefore, few, if any, graphical methods have been explored for multivariate longitudinal data.

Considerable research has been carried out to address covariance modeling of multivariate longitudinal data. These include Kronecker product covariance structures [19,6,18] and random-effects models [24,3,4]. Kronecker product

* Corresponding author.
E-mail address: pkohli@conncoll.edu (P. Kohli).

http://dx.doi.org/10.1016/j.jmva.2015.11.014

structures provide a parsimonious model as the product of the marginal sources of correlations, but their appropriateness in any application depends on whether the underlying conditional independence assumption is tenable. Random-effects models use multivariate mixed models with random coefficients and select the covariance structures from candidate sets like compound symmetry (CS) and autoregressive (AR) models. This approach could lead to model misspecification if the selected covariance structure is far from the truth [31]. For excellent reviews of the current literature from the perspectives of random-effects models and dimension reduction, see [2,28].

For univariate longitudinal data, a statistically interpretable and unconstrained reparameterization of its covariance matrix $\boldsymbol{\Sigma}$ is achieved using its modified Cholesky decomposition: $\boldsymbol{T}\boldsymbol{\Sigma}\boldsymbol{T}^\top = \boldsymbol{D}$, where $\boldsymbol{T}$ is a unit lower triangular matrix with ones on the main diagonal and $\boldsymbol{D}$ is a diagonal matrix with positive diagonal elements. The below-diagonal elements of $\boldsymbol{T}$ are the unconstrained regression coefficients when a measurement is regressed on its predecessors and the diagonal elements of $\boldsymbol{D}$ are the corresponding innovation (regression residual) variances. The non-redundant and unconstrained entries of $\boldsymbol{T}$ and $\ln \boldsymbol{D}$ are then modeled parsimoniously using covariates [22,21]. This idea has been extended to the multivariate longitudinal data setting [15,29] by replacing the scalar entries of $\boldsymbol{T}$ and $\boldsymbol{D}$ by $J \times J$ block matrices

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ -\boldsymbol{\Phi}_{21} & \boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ -\boldsymbol{\Phi}_{31} & -\boldsymbol{\Phi}_{32} & \boldsymbol{I} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\boldsymbol{\Phi}_{T1} & -\boldsymbol{\Phi}_{T2} & \cdots & -\boldsymbol{\Phi}_{T,T-1} & \boldsymbol{I} \end{pmatrix}, \qquad \boldsymbol{D} = \begin{pmatrix} \boldsymbol{D}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_2 & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{D}_T \end{pmatrix}. \tag{1}$$

In this Cholesky *block* decomposition, the blocks in $\boldsymbol{T}$ are unconstrained, interpretable as regression coefficient matrices and they are modeled using established methods reviewed in Section 2.2. However, the diagonal blocks $\boldsymbol{D}_1, \ldots, \boldsymbol{D}_T$ of $\boldsymbol{D}$, which represent the cross-sectional dependence among the regression residuals for the J outcomes, are constraint to be positive-definite. Unlike in the univariate case, a major challenge is finding statistically interpretable and possibly unconstrained reparameterizations for all $\boldsymbol{D}_t$'s simultaneously. To this end, Xu and Mackenzie [29] proposed a log-linear covariance model [7] for the $\boldsymbol{D}_t$'s but the ensuing parameters are difficult to estimate and not interpretable [5], except when $\boldsymbol{D}_t$'s are diagonal. Kim and Zimmerman [15] applied the Cholesky decomposition to each $\boldsymbol{D}_t$, but doing so requires imposing an a priori (time) order among the multiple outcomes which is unnatural in most practical situations.

To address some of the above challenges, we develop a new model for the covariance structure of multivariate longitudinal data. Our approach does not require order among the multiple outcomes and guarantees positive-definiteness of the estimated covariance matrix. It employs various data features to reduce the number of parameters, and motivates a new graphical tool to visualize the multivariate covariance structure. The key idea is to model each $\boldsymbol{D}_t$ using the enhanced *linear covariance models* (LCM):

$$\boldsymbol{D}_t^{\pm 1} = \alpha_1 \boldsymbol{M}_1 + \cdots + \alpha_t \boldsymbol{M}_t, \tag{2}$$

where the $\boldsymbol{M}_i$'s are known positive-definite matrices and the α_i's are unknown non-negative parameters. These restrictions guarantee the positive-definiteness of the postulated covariance models. The enhanced LCM differs from the classic LCM in [1] which only requires the $\boldsymbol{M}_i$'s to be symmetric matrices and α_i's to be scalars. Ensuring the latter has hampered the widespread use of the LCM in practice for over four decades as it involves a computationally demanding procedure for estimating feasible α_i's [25,14,33].

A key issue in LCM is the appropriate choice of $\boldsymbol{M}_i$'s (covariates), an issue which is also present implicitly in the approaches of Xu and Mackenzie [29] and Kim and Zimmerman [15]. We present two methods for selecting and formulating reasonable and positive-definite $\boldsymbol{M}_i$'s. The first method selects the $\boldsymbol{M}_i$'s from a library of known (parametric) covariance basis matrices referred to as *power correlation structures* in [31]. The second method formulates $\boldsymbol{M}_i$'s (nonparametrically) using the eigenvectors of the sample innovation matrix in the spirit of principal component analysis (PCA) as in [12]. For the sake of contrasting the challenges of modeling the covariance of univariate and multivariate longitudinal data, we propose a third method based on extending the notion of univariate regressograms [22] to a multivariate setting. Doing so, unfortunately leads to $\boldsymbol{M}_i$'s that are not necessarily positive-definite.

Once parametric models for the Cholesky factors $(\boldsymbol{T}, \boldsymbol{D})$ of the covariance matrix of the multivariate longitudinal data are identified, the parameters are estimated using the maximum likelihood estimation. For normally distributed data, we solve the more difficult problem of maximum likelihood estimation of the α_i's using a maximization–minimization (MM) procedure [13]. The remainder of this paper is organized as follows. Section 2 describes the development of the modified Cholesky decomposition paired with LCM models for the innovation covariances. Section 3 develops the MM estimation procedure and its asymptotic properties, with an emphasis on how the MM procedure reduces the task of maximizing the likelihood function to minimizing a quadratic function subject to the non-negativity constraint on α_i's in (2). Sections 4 and 5 present numerical results, from Monte Carlo simulations and analysis of data from a bivariate longitudinal study of poplar tree growth. Section 6 concludes the paper. All proofs and derivations are provided in the Supplementary Material (see Appendix A).