Proceedings of the GRASPA 2019 Conference
Pescara, 15-16 July 2019

Edited by: Michela Cameletti, Luigi Ippoliti, Alessio Pollice

Università degli Studi di Bergamo

2019

Proceedings of the GRASPA 2019 Conference
Pescara, 15-16 July 2019

Edited by: Michela Cameletti, Luigi Ippoliti, Alessio Pollice. –

Bergamo : Università degli Studi di Bergamo, 2019.
(GRASPA Working Papers)

ISBN: 978-88-97413-34-9
ISSN: 2037-7738

https://aisberg.unibg.it/handle/10446/142407

Table of Contents

Autoregressive random effects models for circular longitudinal data using the embedding approach

A. Maruotti[1,2,*] and M. Ranalli[3]

[1] *Dipartimento di Giurisprudenza, Economia, Politica e Lingue Moderne. Libera Università Maria Ss Assunta, Via Pompeo Magno 22 - 00192 Roma; a.maruotti@lumsa.it*
[2] *Deparment of Mathematics. University of Bergen; Antonello.Maruotti@uib.no*
[3] *Dipartimento di Economia e Finanza, Università degli Studi di Roma Tor Vergata, Rome, Italy*
[*] *Corresponding author*

Abstract. *Some conditional models to deal with circular longitudinal responses are proposed, extending random effects models to include serial dependence of Markovian form, and hence allowing for quite general association structures between repeated observations recorded on the same unit. The presence of both these components implies a form of dependence between them, and so a complicated expression for the resulting likelihood. To handle this problem, we introduce an approximate conditional mode and a full conditional model, with no assumption about the distribution of the time-varying random effects. All of the discussed models are estimated by means of an EM algorithm for nonparametric maximum likelihood.*

Keywords. *Hidden Markov models; Initial conditions; Finite mixtures; Conditional models.*

1 The embedding approach

Let us introduce the random vector $\mathbf{Y}_{it}$, for unit $i = 1,\dots,I$ at time $t = 1,\dots,T$, following a d-dimensional Normal distribution, with mean $\boldsymbol{\mu}_{it}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e. $\mathbf{Y}_{it} \sim N_d(\boldsymbol{\mu}_{it}, \boldsymbol{\Sigma})$. The random unit vector

$$\mathbf{U}_{it} = \frac{\mathbf{Y}_{it}}{||\mathbf{Y}_{it}||}$$

is said to follow a projected Normal distribution, i.e. $\mathbf{U}_{it} \sim PN_d(\boldsymbol{\mu}_{it}, \boldsymbol{\Sigma})$; see Wang and Gelfand (2013). The general version of the projected normal distribution allows asymmetry and bimodality, i.e different shapes can be modelled. However, the general projected normal distribution is not identified and substantially increases the computational burden required in the estimation step. The distribution of $\mathbf{U}_{it}$ is unchanged if $(\boldsymbol{\mu}_{it}, \boldsymbol{\Sigma})$ is replaced by $(c\boldsymbol{\mu}_{it}, c^2\boldsymbol{\Sigma})$ for any $c > 0$, but this lack of identifiability can be addressed imposing constraints on $\boldsymbol{\Sigma}$. Wang and Gelfand (2013) suggest to set one of the variances in $\boldsymbol{\Sigma}$ to 1 to provide identifiability, resulting in a four-parameter distribution. Other constraints could be also considered as e.g. restricting the determinant of $\boldsymbol{\Sigma}$ to equal 1.

The $\mathbf{U}_{it}$ variable can be converted to an angular random variable, say Θ_{it}, relative to some direction treated as 0. Indeed, any Θ_{it} can be obtained from the radial projection of the bivariate normal distribution by using the *arctan** function defined by Jammalamadaka and SenGupta (2001; p. 13), i.e.

$\Theta_{it} = arctan^* \left(\frac{Y_{it2}}{Y_{it1}}\right) = arctan^* \left(\frac{U_{it2}}{U_{it1}}\right)$. The following explicit relation exists between $\mathbf{Y}_{it}$ and the circular variable Θ_{it}

$$\mathbf{Y}_{it} = \begin{bmatrix} Y_{it1} \\ Y_{it2} \end{bmatrix} = \begin{bmatrix} R_{it}\cos\theta_{it} \\ R_{it}\sin\theta_{it} \end{bmatrix} = R_{it}\mathbf{U}_{it},$$

where $R_{it} = ||\mathbf{Y}_{it}||$.

In the following, we will focus exclusively on the case $\mathbf{\Sigma} = \mathbf{I}$ and $d = 2$ (i.e. on circular data). If in addition, $\boldsymbol{\mu}_{it} = \mathbf{0}$, then $\mathbf{U}_{it}$ is uniformly distributed on the circle; otherwise the distribution of $\mathbf{U}_{it}$ is unimodal and rotationally symmetric about its mean direction $\boldsymbol{\mu}_{it}/||\boldsymbol{\mu}_{it}||$. Indeed, departure from zero for the two means, in the case of an identity covariance matrix, creates one mode in the trigonometric quadrant with the same sign of the means, e.g. if $\mu_{it1} > 0$ and $\mu_{it2} < 0$, where $\boldsymbol{\mu}_{it} = (\mu_{it1}, \mu_{it2})$, then the mode is in the quadrant with positive cosine and negative sine.

The joint density $f(\theta_{it}, r_{it} \mid \boldsymbol{\mu}_{it}, \mathbf{I})$ can be easily obtained by transforming the bivariate normal distribution of $\mathbf{y}_{it}$ to polar coordinates, i.e. $f(\theta_{it}, r_{it} \mid \boldsymbol{\mu}_{it}, \mathbf{I}) = f(r_{it}\mathbf{u}_{it} \mid \boldsymbol{\mu}_{it}, \mathbf{I})r_{it}$ and, thus, $f(\theta_{it} \mid \boldsymbol{\mu}_{it}, \mathbf{I}) = \int f(\theta_{it}, r_{it} \mid \boldsymbol{\mu}_{it}, \mathbf{I})dr_{it} = \phi(\mu_{it1}, \mu_{it2}; \mathbf{0}, \mathbf{I}) + \mu_{it1}\cos\theta_{it} + \mu_{it2}\sin\theta_{it}$, i.e. $\theta_{it} \sim PN_2(\boldsymbol{\mu}_{it}, \mathbf{I})$, with $\phi(\cdot)$ denoting the density function of the bivariate normal distribution.

In empirical applications, the angle θ_{it} is usually collected. However, as discussed above, we prefer to work with its radial projections as the resulting model can be easily dealt with by using standard regression modelling strategies. In other words, we model $\mathbf{Y}_{it}$ and focus our interest upon the parameter vector $\boldsymbol{\mu}_{it}$, which is modelled, in a regression framework, by defining a multivariate linear mixed model, as defined in the following sections.

2 The random effects model

The temporal evolution of the random effects can be conveniently described by including a vector of time-varying random effects, say $\mathbf{b}_{it} = (b_{it1}, b_{it2})$. Regarding $\mathbf{b}_{it}$'s distribution, we assume a (hidden) Markov chain with states $\mathbf{b}_k = (b_{k1}, b_{k2}), k = 1, \ldots, K$, initial probabilities $\pi_{ik} = \Pr(\mathbf{b}_{i1} = \mathbf{b}_k) = \pi_k$ and transition probability matrix $\mathbf{\Pi} = \{\pi_{it,k|h}\}$ with $\pi_{it,k|h} = \Pr(\mathbf{b}_{it} = \mathbf{b}_k \mid \mathbf{b}_{it-1} = \mathbf{b}_h) = \pi_{k|h}, t > 1$, i.e. Markov chain's parameters will be assumed independent on any covariates and shared among subjects.

The modelling framework is completed by defining the regression model (see also Maruotti et al., 2016)

$$\mu_{itj} = \mathbf{x}'_{it}\boldsymbol{\beta}_j + b_{itj}, \quad j = 1, 2,$$

where $\mathbf{x}_{it} = (1, x_{it1}, \ldots, x_{itp}, \theta_{it-1})$, $\boldsymbol{\beta}_j = (\beta_{0j}, \beta_{1j}, \ldots, \beta_{pj}, \beta_{p+1,j})$ represents the $(p+2)$-dimensional vector of regression parameters referred to the j-th projection and $\mathbf{b}_i = (b_{i1}, b_{i2})$ denotes a set of subject- and projection-specific random effects. However, in order to easily implement the estimation steps, the following multilevel specification can be considered

$$\mu_{itj} = \sum_{j=1}^{2} d_{itj}(\mathbf{x}'_{it}\boldsymbol{\beta}_j + b_{itj})$$

where we use a set of indicator variables d_{itj}, with $d_{itj} = 1, \forall i = 1, \ldots, I; t = 1, \ldots, T, j = 1, 2$ iff the j-th projection is to be modelled and 0 otherwise. Using a matrix notation

$$\boldsymbol{\mu}_{it} = \mathbf{x}^*_{it}\boldsymbol{\beta}^* + \mathbf{b}^*_{it}$$

where
$$\boldsymbol{\mu}_{it} = \begin{bmatrix}\mu_{it1}\\ \mu_{it2}\end{bmatrix}, \quad \mathbf{x}^*_{it} = \mathbf{I}_2 \otimes \mathbf{x}_{it}, \quad \boldsymbol{\beta}^* = vec(\boldsymbol{\beta}) = \begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{bmatrix}, \quad \mathbf{b}^*_{it} = vec(\mathbf{b}_{it}) = \begin{bmatrix}b_{it1}\\ b_{it2}\end{bmatrix}.$$

If the covariates are not the same for both projections, some of the elements of $\boldsymbol{\beta}_j$ would be set equal to zero.

We would remark that the circular mean direction and concentration, i.e., the circular counterpart of the mean and precision of a linear random variable, are respectively $\bar{\mu}_{it} = arctan^*\left(\frac{\mu_{it2}}{\mu_{it1}}\right)$ and $c_{it} = (\pi\gamma_{it}/2)^{1/2}\exp(-\gamma_{it})(I_0(\gamma_{it}) + I_1(\gamma_{it}))$, where $\gamma_{it} = ||\boldsymbol{\mu}_{it}||^2/4$ and $I_v(\gamma)$ is the modified Bessel function of the first kind of order v, see Wang and Gelfand (2013). Both $\bar{\mu}_{it}$ and c_{it} depend on the means of the projections hence the regression type specification of μ_{itj} can adjust for change in mean direction and concentration due to different levels of covariates.

2.1 Likelihood inference

Inference for the proposed model is based on the log-likelihood
$$\ell(\boldsymbol{\lambda}) = \sum_{i=1}^{I} \log\left\{\sum_{\mathbf{b}_{i1}}\cdots\sum_{\mathbf{b}_{iT}}\left[\pi_{\mathbf{b}_{i1}}\prod_{t>1}\pi_{\mathbf{b}_{it}|\mathbf{b}_{it-1}}\prod_{t} f(\theta_{it} \mid \mathbf{x}_{it}, \mathbf{b}_{it}) f(\theta_{i0} \mid \mathbf{x}_{i0}, \mathbf{b}_{i0})\right]\right\}$$
with the sum $\sum_{\mathbf{b}_{it}}$ extended to all possible configurations of $\mathbf{b}_{it}$ and where $\boldsymbol{\lambda}$ is a short-hand notation for all non-redundant parameters. However, inferences can be highly sensitive to misspecification of $f(\theta_{i0} \mid \mathbf{x}_{i0}, \mathbf{b}_{i0})$. Thus, we rewrite the previous expression as
$$\ell(\boldsymbol{\lambda}) = \sum_{i=1}^{I} \log\left\{\sum_{\mathbf{b}_{i1}}\cdots\sum_{\mathbf{b}_{iT}}\left[\pi_{\mathbf{b}_{i1}}(\theta_{i0})\prod_{t>1}\pi_{\mathbf{b}_{it}|\mathbf{b}_{it-1}}(\theta_{i0})\prod_{t} f(\theta_{it} \mid \mathbf{x}_{it}, \mathbf{b}_{it}) f(\theta_{i0} \mid \mathbf{x}_{i0})\right]\right\}$$
or equivalently
$$\ell(\boldsymbol{\lambda} \mid \theta_{i0}) = \sum_{i=1}^{I} \log\left\{\sum_{\mathbf{b}_{i1}}\cdots\sum_{\mathbf{b}_{iT}}\left[\pi_{\mathbf{b}_{i1}}(\theta_{i0})\prod_{t>1}\pi_{\mathbf{b}_{it}|\mathbf{b}_{it-1}}(\theta_{i0})\prod_{t} f(\theta_{it} \mid \mathbf{x}_{it}, \mathbf{b}_{it})\right]\right\}$$
resulting in a conditional likelihood, where $\pi_{\mathbf{b}_{i1}}(\theta_{i0})$ and $\pi_{\mathbf{b}_{it}|\mathbf{b}_{it-1}}(\theta_{i0})$ allows the random effects distribution to dependent on θ_{i0}.

References

[1] Jammalamadaka RA, SenGupta A (2001). *Topics in Circular Statistics*. World Scientific

[2] Maruotti A, Punzo A, Mastrantonio G and Lagona F (2016). A time-dependent extension of the projected normal regression model for longitudinal circular data based on a hidden Markov heterogeneity structure. *Stochastic Environmental Research and Risk Assessment*, 30: 1725-1740.

[3] Wang F, Gelfand A (2013). Directional data analysis under the general projected normal distribution. *Statistical Methodology* **10**: 113–127.