Nonparametric Covariance Estimation for Longitudinal Data via Penalized Tensor Product Splines

Tayler A. Blake* Yoonkyung Lee†

January 9, 2018

0.1 Penalized likelihood estimation

Let Y hold the N observed response vectors $y_1, \ldots, y_N$ less their first element y_{i1} stacked into a single vector of dimension $n_y = \left(\sum_i M_i \right) - N$. Let M denote the total number of distinct observation times across all subjects. For ease of exposition, let $\sigma_{ij} = \sigma\left(t_{ij}\right)$ and $\phi_{ijk} = \phi\left(t_{ijk}\right)$. The loglikelihood **??** becomes

$$
\begin{aligned}
-2\ell\left(Y, \Sigma\right) &= \sum_{t=1}^{M} \log \sigma_t^2 + \sum_{i=1}^{N} \sum_{j=1}^{m_i} \frac{\epsilon_{ijk}^2}{\sigma_{ij}^2} \\
&= \sum_{t=1}^{M} \log \sigma_t^2 + \sum_{i=1}^{N} \frac{\epsilon_{i1}^2}{\sigma_{i1}^2} + \sum_{i=1}^{N} \sum_{j=2}^{m_i} \frac{\epsilon_{ij}^2}{\sigma_{ij}^2} \\
&= \sum_{t=1}^{M} \log \sigma_t^2 + \sum_{i=1}^{N} \frac{y_{i1}^2}{\sigma_{i1}^2} + \sum_{i=1}^{N} \sum_{j=2}^{m_i} \sigma_{ij}^{-2} \left(y_{ij} - \sum_{k<j} \phi_{ijk} y_{ik} \right)^2 .
\end{aligned} \tag{1}
$$

$$
\sum_{t=1}^{M} \log \sigma_t^2 + \sum_{i=1}^{N} \frac{y_{i1}^2}{\sigma_{i1}^2} + \sum_{i=1}^{N} \sum_{j=2}^{m_i} \frac{\epsilon_{ij}^2}{\sigma_{ij}^2} \tag{2}
$$

An iterative procedure for minimizing **??** starts by first initialising ?t, using for example the innovation standard error estimated without the penalty. We then minimise (10) to obtain ?tj, j = 1,...,t?1, and revise ?t2 as in (9). We iterate the process until convergence for each t, t = 2, , n. For details about minimisation of (10) with fixed ?t see the Appendix.

*The Ohio State University, 1958 Neil Avenue, Columbus, OH 43201
†The Ohio State University, 1958 Neil Avenue, Columbus, OH 43201

1 Computation of the smoothing spline estimator

2 Computation of the P-spline estimator

3 Smoothing parameter selection