Nonparametric covariance estimation for longitudinal data via tensor product smoothing

Tayler A. Blake

June 27, 2017

The data:

$$Y_i = (Y_{i1}, Y_{i2}, \ldots, Y_{im})', \qquad i = 1, \ldots, N$$

associated with measurement times

$$t_1 < t_2 < \cdots < t_m.$$

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite.
  - Constrained optimization is a headache.
- The $\{t_{ij}\}$ may be suboptimal.
  - Observation times may not fall on a regular grid, may vary across subjects.
- More dimensions, more problems (maybe.)
  - Sample covariance matrix falls apart when m is large.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite.
  - Constrained optimization is a headache.
- The $\{t_{ij}\}$ may be suboptimal.
  - Observation times may not fall on a regular grid, may vary across subjects.
- More dimensions, more problems (maybe.)
  - Sample covariance matrix falls apart when m is large.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite.
  - Constrained optimization is a headache.
- The $\{t_{ij}\}$ may be suboptimal.
  - Observation times may not fall on a regular grid, may vary across subjects.
- More dimensions, more problems (maybe.)
  - Sample covariance matrix falls apart when m is large.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite. A cute little reparameterization $\Longrightarrow$ unconstrained estimation, meaningful interpretation
- The $\{t_{ij}\}$ may be messy.
  Frame covariance estimation as function estimation.
- More dimensions, more problems (maybe.)

Figure: Regulate like Nate Dogg.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite. A cute little reparameterization $\implies$ unconstrained estimation, meaningful interpretation
- The $\{t_{ij}\}$ may be messy.
  Frame covariance estimation as function estimation.
- More dimensions, more problems (maybe.)

Figure: Regulate like Nate Dogg.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite. A cute little reparameterization $\Longrightarrow$ unconstrained estimation, meaningful interpretation
- The $\{t_{ij}\}$ may be messy.
  Frame covariance estimation as function estimation.
- More dimensions, more problems (maybe.)

Figure: Regulate like Nate Dogg.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite. A cute little reparameterization $\Longrightarrow$ unconstrained estimation, meaningful interpretation
- The $\{t_{ij}\}$ may be messy.
  Frame covariance estimation as function estimation.
- More dimensions, more problems (maybe.)

Figure: Regulate like Nate Dogg.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite. A cute little reparameterization $\implies$ unconstrained estimation, meaningful interpretation
- The $\{t_{ij}\}$ may be messy.
  Frame covariance estimation as function estimation.
- More dimensions, more problems (maybe.)

Figure: Regulate like Nate Dogg.

The flaming hoops:

- Covariance matrices (and their estimates) should be positive definite. A cute little reparameterization $\implies$ unconstrained estimation, meaningful interpretation
- The $\{t_{ij}\}$ may be messy.
  Frame covariance estimation as function estimation.
- More dimensions, more problems (maybe.)

Figure: Regulate like Nate Dogg.

Covariance dress-up: the modified Cholesky decomposition

$$Y = (Y_1, \ldots, Y_M)' \sim \mathcal{N}(0, \Sigma).$$

For any positive definite Σ, we can find T which diagonalizes Σ:

$$D = T\Sigma T', \quad T = \begin{bmatrix} 1 & 0 & \ldots & & \\ -\phi_{21} & 1 & & & \\ -\phi_{31} & -\phi_{32} & 1 & & \\ \vdots & & & \ddots & \\ -\phi_{M1} & -\phi_{M2} & \ldots & -\phi_{M,M-1} & 1 \end{bmatrix} \tag{1}$$

Here's the cute part:

Imagine regressing Y_j on its predecessors:

$$Y_j = \begin{cases} e_1 & j = 1, \\ \sum_{k=1}^{j-1} \phi_{jk} Y_k + \sigma_j e_j & j = 2, \ldots, M \end{cases} \tag{2}$$

In matrix form:

$$e = TY, \tag{3}$$

and taking covariances on both sides:

$$D = diag\left(\sigma_1^2, \ldots, \sigma_M^2\right) = T\Sigma T'. \tag{4}$$

No constraints on the ϕ_{jk}s!

The regression model tool box: a deep treasure chest of luxury.

Model Y_j, e_j as

$$Y_j = Y\left(t_j\right), \quad e_j = e\left(t_j\right),$$
$$e\left(s\right) \sim \mathcal{WN}\left(0,1\right),$$

Swap the standard regression model 2 for a varying coefficient model:

$$\phi_{jk} = \phi\left(t_j, t_k\right),$$

$$y\left(t_j\right) = \sum_{k=1}^{j-1} \phi\left(t_j, t_k\right) y\left(t_k\right) + \sigma\left(t_j\right) \epsilon\left(t_j\right) \tag{5}$$

(Iterated) penalized maximum likelihood estimation

1. Fix $\sigma_{ij}^2 = \sigma_{ij0}^2$, $i = 1, \dots, N$,$j = 1, \dots, M$.
2. Find $\phi_0 = \underset{\phi}{arg\ min} - 2L_\phi\left(\phi, y_1, \dots, y_N\right) + \lambda J\left(\phi\right)$
3. Fix $\phi = \phi_0$.
4. Find $\sigma_0^2 = \underset{\sigma^2}{arg\ min} - 2L_\sigma^2\left(\sigma^2, y_1, \dots, y_N\right) + \lambda J\left(\sigma^2\right)$

$$-2L_\phi\left(\phi, y_1, \dots, y_N\right) = \sum_{i=1}^{N}\sum_{j=2}^{m_i} \sigma_{ij0}^{-2}\left(y_{ij} - \sum_{k=1}^{j-1}\phi\left(t_{ij}, t_{ik}\right) y_{ik}\right)^2$$

Regularization of $\phi(s,t)$ is more intuitive if we transform the s-t axis.

Rotate the input axes:

$$
\begin{aligned}
l &= s - t \\
m &= \frac{1}{2}(s+t).
\end{aligned}
$$

Then ϕ becomes

$$
\begin{aligned}
\phi^*(l,m) &= \phi^*\left(s-t, \frac{1}{2}(s+t)\right) \\
&= \phi(s,t).
\end{aligned}
$$

Take $\hat{\phi}^*$ to be the minimizer of

$$-2L + \lambda J(\phi^*)$$

Smooth ANOVA models

Decompose

$$\phi^*(l,m) = \mu + \phi_1(l) + \phi_2(m) + \phi_{12}(l,m), \qquad (7)$$

so Model 5 becomes

$$y(t_j) = \sum_{k=1}^{j-1} \Big[\mu + \phi_1(l_{jk}) + \phi_2(m_{jk}) \\ + \phi_{12}(l_{jk}, m_{jk}) \Big] y(t_k) + \sigma(t_j)\,\epsilon(t_j)$$