Nonparametric covariance estimation for longitudinal data via tensor product smoothing

Tayler Blake [1]   Dr. Yoonkyung Lee [2]

[1]Information Control Company

[2]The Ohio State University, Department of Statistics

July 8, 2017

The data:

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

associated with measurement times

$$t_1 < t_2 < \cdots < t_{M_i}.$$

Goal: estimate

$$Cov\,(Y) = \Sigma$$

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 $\implies$ 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:

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

Now, for the cutest part:

Okay, really:

Regress Y_j on $Y_{(1:j-1)} = (Y_1, \ldots, Y_{j-1})'$:

$$y_j = \begin{cases} e_1 & j = 1, \\ \sum_{k=1}^{j-1} \phi_{jk} y_k + \sigma_j e_j & j = 2, \ldots, M \end{cases} \tag{1}$$
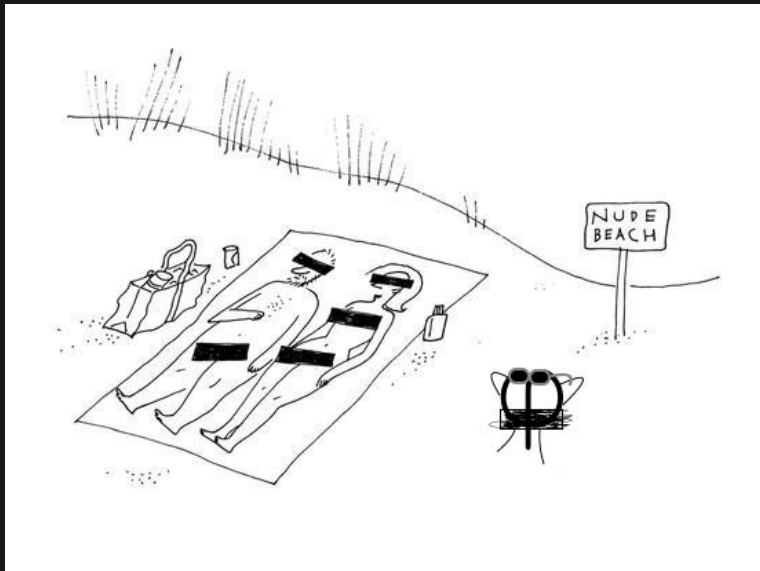
In matrix form:

$$e = TY, \tag{2}$$

and taking covariances on both sides:

$$D = diag\left(\sigma_1^2, \ldots, \sigma_M^2\right) = T\Sigma T'. \tag{3}$$

The regression model tool box is a deep, luxurious toolbox.

$$Y_j \longrightarrow Y(t_j) \qquad\qquad e_j \longrightarrow e(t_j)$$
$$\phi_{jk} \longrightarrow \phi(t_j, t_k) \qquad\qquad \sigma_j^2 \longrightarrow \sigma^2(t_j)$$

$$y(t_j) = \sum_{k=1}^{j-1} \phi(t_j, t_k) \, y(t_k) + \sigma(t_j) \, e(t_j), \qquad (4)$$

where

$$e(s) \sim \mathcal{WN}(0, 1)$$

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

$$l = s - t$$
$$m = \frac{1}{2}(s + t)$$

Reparameterize ϕ:

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

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

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

Smooth ANOVA models

Decompose

$$\phi^*(l, m) = \mu + \phi_1(l) + \phi_2(m) + \phi_{12}(l, m), \qquad (6)$$

so Model 4 becomes

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

Approximate $\phi_1,\ \phi_2,\ \phi_{12}$ with B-splines.

$$\phi_1 (l) = \sum_{c=1}^{c_l} B_c (l; q_l) \theta_{lc} = B_l \theta_l,$$

$$\phi_2 (m) = \sum_{c'=1}^{c_m} B_{c'} (m; q_m) \theta_{mc'} = B_m \theta_m \qquad (7)$$

$$\phi_{12} (l, m) = \sum_{c=1}^{c_l} \sum_{c'=1}^{c_m} B_c (l; q_l) B_{c'} (m; q_m) \theta_{cc'} = B_{lm} \theta_{lm}$$

$$B_{lm} = B_m \,\square\, B_l$$
$$\equiv \left(B_m \otimes 1'_{c_l} \right) \odot \left(1'_{c_m} \otimes B_l \right)$$

13

Difference penalty had to regulate.

For $f(x) = \sum\limits_{i=1}^{p} B_i(x)\theta_i$, approximate

$$\int_0^1 (f''(x))^2 \; dx = \int_0^1 \left\{ \sum_{i=1}^{p} B_i''(x)\theta_i \right\}^2 dx \tag{8}$$
$$= k_1 \sum_i (\Delta^2\theta_i)^2 + k_2,$$

by

$$||D_2\theta||^2, \qquad D_2\theta = (\Delta^2\theta_1, \ldots, \Delta^2\theta_{p-2})'$$

In general, $\boxed{\text{approximate } \int\limits_0^1 \left(f^{(d)}\right)^2 dx \text{ with } ||D_d\theta||^2}$

PS-ANOVA Penalty

$$B \equiv [\, 1_p \mid B_l \mid B_m \mid B_{lm} \,], \qquad \theta \equiv (\theta_l, \theta_m, \theta_{lm})'$$

Find θ minimizing

$$(Y - WB\theta)' (Y - WB\theta) + \theta' P \theta$$

$$P = \begin{bmatrix} \underbrace{\lambda_l D'_{d_l} D_{d_l}}_{P_l} & & \\ & \underbrace{\lambda_l D'_{d_l} D_{d_l}}_{P_m} & \\ & & \underbrace{\tau_l D'_{d_l} D_{d_l} \otimes I_{c_m} + \tau_m I_{c_l} \otimes D'_{d_m} D_{d_m}}_{P_{lm}} \end{bmatrix}$$

PS-ANOVA Penalty

$$B \equiv [\, 1_p \mid B_l \mid B_m \mid B_{lm} \,], \qquad \theta \equiv (\theta_l, \theta_m, \theta_{lm})'$$

Find θ minimizing

$$(Y - WB\theta)'\, D^{-1}\, (Y - WB\theta) + \theta' P \theta \tag{9}$$

$$P = \begin{bmatrix} \underbrace{\lambda_l D'_{d_l} D_{d_l}}_{P_l} & & \\[2em] & \underbrace{\lambda_l D'_{d_l} D_{d_l}}_{P_m} & \\[2em] & & \underbrace{\tau_l D'_{d_l} D_{d_l} \otimes I_{c_m} + \tau_m I_{c_l} \otimes D'_{d_m} D_{d_m}}_{P_{lm}} \end{bmatrix}$$

Mixed model representation

Find transformation $Q = \begin{bmatrix} Q_n \mid Q_s \end{bmatrix}$ to map

$$BQ = \begin{bmatrix} BQ_n \mid BQ_s \end{bmatrix} \qquad Q'\theta = (\beta', \alpha')'$$
$$= \begin{bmatrix} X \mid Z \end{bmatrix} \qquad\qquad = \left((Q_n'\theta)', \ (Q_s'\theta)'\right)'$$

to reparameterize the ill-posed Model 9 as

$$Y = W(X\beta + Z\alpha) + e$$
$$\alpha \sim \mathcal{N}(0, G), \qquad\qquad (10)$$
$$e \sim \mathcal{N}(0, D)$$

Mixed model representation

$$
\begin{bmatrix}
\lambda_l \tilde{\Delta}_l & & \\
& \lambda_m \tilde{\Delta}_m & \\
& & \begin{bmatrix} \tau_l \tilde{\Delta}_l & & \\ & \tau_m \tilde{\Delta}_m & \\ & & \tau_m \tilde{\Delta}_m \otimes I_{c_l - d_l} + I_{c_m - d_m} \otimes \tau_l \tilde{\Delta}_l \end{bmatrix}
\end{bmatrix}
$$

Decomposition of ϕ^* for $d_l = d_m = 2$

	$\{1\}$	$\{m\}$	$\{B_{j'}(m)\}$
$\{1\}$	$\{1\}$	$\{m\}$	$\{B_{j'}(m)\}$
$\{l\}$	$\{l\}$	$l \times m$	$l \times \{B_{j'}(m)\}$
$\{B_j(l)\}$	$\{B_j(l)\}$	$m \times \{B_j(l)\}$	$\{B_j(l)\,B_{j'}(m)\}$

Nested PS-ANOVA

Re-express ϕ_{12}:

$$\phi_{12}\left(l, m\right) = g_1\left(l\right)\Big[\sum_{r=1}^{d_m-1} m^r\Big] + \Big[\sum_{r=1}^{d_l-1} l^r\Big]g_2\left(m\right) + h\left(l, m\right),$$

For $d_l = d_m = 2$,

$$\phi_{12}\left(l, m\right) = g_1\left(l\right)\ m + l\ g_2\left(m\right) + h\left(l, m\right)$$

with basis:

$$B = \big[\ 1_p\ |\ B_1\ |\ B_2\ |B_3\ |B_4\ |B_5\ \big], \tag{11}$$

where

$$B_3 = m\ \square\ B_1 \qquad\qquad B_5 = B_2\ \square\ B_1$$
$$B_4 = B_2\ \square\ l$$

Nested PS-ANOVA

$$P = \text{blockdiag}\left(0,\ P_1,\ P_2,\ P_3,\ P_4,\ P_5\right), \qquad (12)$$

Nested PS-ANOVA

Re-express ϕ_{12}:

$$\phi_{12}(l, m) = g_{lr}(l) \Big[\sum_{r=1}^{d_m - 1} m^r \Big] + \Big[\sum_{r=1}^{d_l - 1} l^{r'} \Big] g_{mr'}(m) + h(l, m),$$

For $d_l = d_m = 2$,

$$\phi_{12}(l, m) = g_{l_1}(l)\ m + l\ g_{m_1}(m) + h(l, m)$$

with basis:

$$B = \big[\ 1_p\ |\ B_1\ |\ B_2\ |B_3\ |B_4\ |B_5\ \big], \qquad (13)$$

where

$$B_3 =\ m \,\square\, B_1 \qquad\qquad B_5 =\ B_2 \,\square\, B_1$$
$$B_4 =\ B_2 \,\square\, l$$

$$P = \text{blockdiag}\,(0,\ P_1,\ P_2,\ P_3,\ P_4,\ P_5)$$

Remove the redundant columns, and give each penalized component its own random effect:

$$B$$