Nonparametric Covariance Estimation for Longitudinal Data

Tayler A. Blake

The Ohio State University, Department of Statistics

April 27, 2018

Background

Let

$$Y = (y_1, \ldots, y_p)', \quad (t_1, \ldots, t_p)'$$

denote the random vector of observations and their associated measurement times, where

$$Cov\,(Y) = \Sigma = [\sigma_{ij}]$$

- Dimensionality: the number of parameters σ_{ij} is quadratic in p.

Background

Let

$$Y = (y_1, \ldots, y_p)', \quad (t_1, \ldots, t_p)'$$

denote the random vector of observations and their associated measurement times, where

$$Cov\,(Y) = \Sigma = [\sigma_{ij}]$$

- Dimensionality: the number of parameters σ_{ij} is quadratic in p.
- The positive definite constraint:

$$c'\Sigma c = \sum_{i,j=1}^{p} c_i c_j \sigma_{ij} \geq 0$$

Background

Let

$$Y = (y_1, \ldots, y_p)', \quad (t_1, \ldots, t_p)'$$

denote the random vector of observations and their associated measurement times, where

$$Cov\,(Y) = \Sigma = [\sigma_{ij}]$$

- Dimensionality: the number of parameters σ_{ij} is quadratic in p.
- The positive definite constraint:

$$c'\Sigma c = \sum_{i,j=1}^{p} c_i c_j \sigma_{ij} \geq 0$$

- Observation times may be irregular or subject-specific.

Popular approaches to covariance modeling

Parametric Models

- autoregressive models
- moving average models
- compound symmetry

Low variance, potentially high bias

- parsimonious, stable
- computationally convenient
- bias from model misspecification

Sample covariance S

$$S = (N-1)^{-1} \sum_{i=1}^{N} \left(Y_i - \bar{Y}\right) \left(Y_i - \bar{Y}\right)'$$

Unbiased for Σ, potentially high variance

- flexible, unbiased
- positive definite
- unstable even when $N > p$ and p is moderate.

Outline

- Background: A Review of Covariance Estimation for Longitudinal Data
  - Parametric Covariance Models
  - Shrinkage Estimators based on the Sample Covariance Matrix
  - Matrix Decompositions and GLMs for Covariance
- A Reproducing Kernel Hilbert Space Framework for Covariance Estimation
- Tensor Product P-splines for Covariance Estimation
- Simulation Studies
- Data Analysis: Cattle Weights

Applying Elementwise Shrinkage to S

Tapering Estimators

- **The Banded Sample Covariance Matrix**

$$B_k(S) = \left[s_{ij} 1 \left(|i-j| \leq k\right)\right] = R_B * S, \quad 0 < k < p.$$

- **The Tapered Sample Covariance Matrix**

$$S^{\omega} = \left[\omega_{ij}^{k} s_{ij}\right],$$

where $0 < k < p$, and if $k_h = k/2$

$$\omega_{ij}^{k} = k_h^{-1}\left[(k - |i-j|)_+ - (k_h - |i-j|)_+\right].$$

- **Soft Thresholding Estimator:**

$$S^{\lambda} = \left[\text{sign}\left(s_{ij}\right)\left(s_{ij} - \lambda\right)_+\right],$$

▸▸ Simulations

The Modified Cholesky decomposition

For any positive definite Σ, there exists a unique lower-triangular matrix $C = [c_{ij}]$, $c_{ii} > 0$:

$$\Sigma = CC',$$

Let $D^{1/2} = diag\,(c_{11}, \ldots, c_{pp})$, $L = D^{-1/2}C$, then

$$\Sigma = LDL'.$$

The **modified Cholesky decomposition** (MCD) of Σ is given by

$$D = T\Sigma T', \tag{1}$$

where $T = L^{-1}$. The lower triangular entries of T are *unconstrained*.

Statistical Interpretation of (T, D)

Let $\hat{y}_t$ be the linear least-squares predictor of y_t based on previous measurements $y_{t-1}, \ldots, y_1$. We can find unique scalars ϕ_{tj}:

$$y_t = \begin{cases} \epsilon_t, & t = 1 \\ \sum_{j=1}^{t-1} \phi_{tj} y_j + \epsilon_t, & t = 2, \ldots, p, \end{cases} \tag{2}$$

where $E[\epsilon_t] = 0$, $D = Cov(\epsilon) = diag\left(\sigma_1^2, \ldots, \sigma_p^2,\right)$. Then

$$\underbrace{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{p-1} \\ \epsilon_p \end{bmatrix}}_{\epsilon} = \underbrace{\begin{bmatrix} 1 & & & & \\ -\phi_{21} & 1 & & & \\ -\phi_{31} & -\phi_{32} & 1 & & \\ \vdots & & & \ddots & \\ -\phi_{p1} & -\phi_{p2} & \ldots & -\phi_{p,p-1} & 1 \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{p-1} \\ y_p \end{bmatrix}}_{Y}$$

Taking the covariance on both sides gives the MCD (1).

The coefficients and prediction error variances of successive regressions are unconstrained.

The **generalized autoregressive parameters** ϕ_{tj} and **log innovation variances** $\log \sigma_j^2$ are unconstrained but

$$\hat{\Sigma}^{-1} = \hat{T}'\hat{D}^{-1}T$$

is guaranteed to be positive definite.

$$\begin{array}{cccccc}
y_1 & y_2 & y_3 & \cdots & y_{p-1} & y_p \\
\hline
1 & & & & & \\
\phi_{21} & 1 & & & & \\
\phi_{31} & \phi_{32} & 1 & & & \\
\vdots & \vdots & & \ddots & & \\
\vdots & \vdots & & & \ddots & \\
\phi_{p1} & \phi_{p2} & \cdots & \cdots & \phi_{p,p-1} & 1 \\
\hline
\sigma_1^2 & \sigma_2^2 & \cdots & \cdots & \sigma_{p-1}^2 & \sigma_p^2
\end{array}$$

Maximum Normal Likelihood Estimation for the Cholesky Decomposition

The MLE for (T, D) has closed form.

For $Y_1, \ldots, Y_N \sim N\left(0_p, \Sigma\right)$ and $S = N^{-1} \sum_{i=1}^{N} Y_i Y_i'$,

$$
\begin{aligned}
-2\ell\left(\Sigma | Y_1, \ldots, Y_N\right) &= \sum_{i=1}^{N} \left(\log|\Sigma| + Y_i' \Sigma^{-1} Y_i'\right) \\
&= N \log|D| + N \mathrm{tr}\left(D^{-1} T S T'\right)
\end{aligned}
\tag{3}
$$

is quadratic in T for fixed D, so the MLE for the ϕ_{tj} has closed form. Similarly, the MLE for D for fixed T has closed form.

Parametric Models for the Cholesky Decomposition

Pourahmadi (2000), Pan andMackenzie (2003) suggest modeling ϕ_{tj}, σ_t^2 with covariates, letting

$$\phi_{jk} = z'_{jk}\gamma$$
$$\log \sigma_j^2 = z'_j\lambda.$$

Common choices for the covariates x_{jk} and z_j are

$$x'_{jk} = \left(1, t_j - t_k, (t_j - t_k)^2, \ldots, (t_j - t_k)^{p-1}\right)', \qquad (4)$$
$$z'_j = \left(1, t_j, \ldots, t_j^{q-1}\right)'.$$

Polynomial orders p and q are tuning parameters chosen by a model selection criterion (BIC, AIC).

Simulations

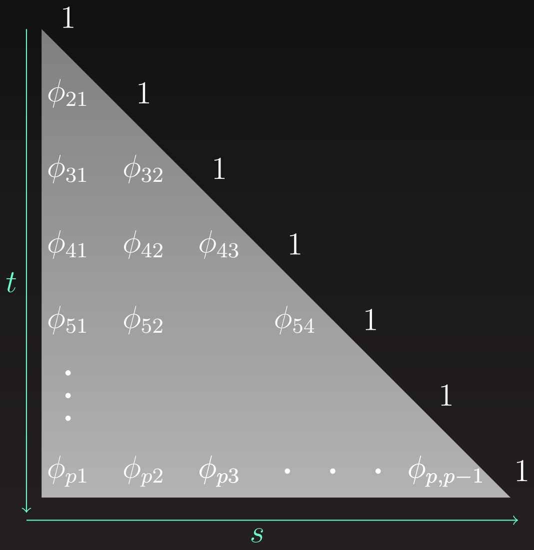
Accommodating Unbalanced Data

by treating ϕ as a continuous bivariate function.

$$
\begin{array}{ccccccc}
1 & & & & & & \\
\phi_{21} & 1 & & & & & \\
\phi_{31} & \phi_{32} & 1 & & & & \\
\phi_{41} & \phi_{42} & \phi_{43} & 1 & & & \\
\phi_{51} & \phi_{52} & & \phi_{54} & 1 & & \\
\vdots & & & & & 1 & \\
\phi_{p1} & \phi_{p2} & \phi_{p3} & \cdots & & \phi_{p,p-1} & 1
\end{array}
$$

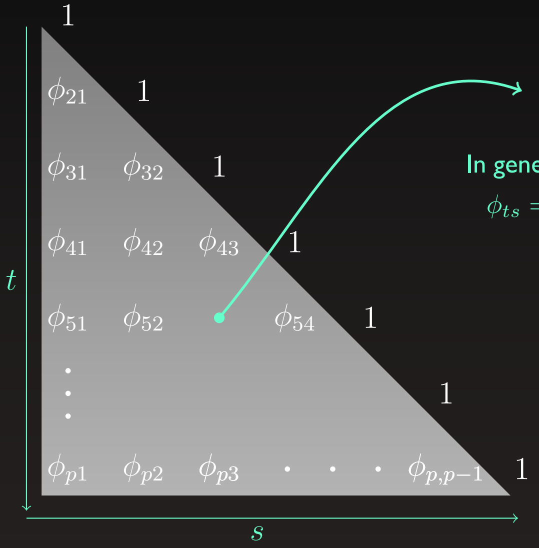
Accommodating Unbalanced Data

by treating ϕ as a continuous bivariate function.



Accommodating Unbalanced Data

by treating ϕ as a continuous bivariate function.



$$\phi_{53} = \tilde{\phi}\left(t_5, t_3\right)$$

In general,

$$\phi_{ts} = \tilde{\phi}\left(t, s\right), \quad 0 \leq s < t \leq 1$$

A functional varying coefficient model for ϕ

Assume measurements $Y_i = (y_{i1}, \ldots, y_{i,p_i})'$ arise from $Y(t)$ observed at

$$t_i = \{t_{i1} < \cdots < t_{i,p_i}\} \subset \mathcal{T} = [0,1]$$

Model

$$\boxed{y(t_{ij}) = \sum_{k<j} \tilde{\phi}(t_{ij}, t_{ik})\, y(t_{ik}) + \epsilon(t_{ij}), \qquad \begin{array}{l} i = 1, \ldots, N \\ j = 1, \ldots, p_i, \end{array}} \tag{5}$$

Transform $l = t - s, \quad m = \frac{t+s}{2}$, let

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

so that

$$-2\ell\left(\phi, \sigma^2 | Y_1, \ldots, Y_N\right) = \sum_{i=1}^{N}\sum_{j=2}^{p_i} \log \sigma_{ij}^2 + \sum_{i=1}^{N}\sum_{j=2}^{p_i} \frac{1}{\sigma_{ij}^2}\left(y_{ij} - \sum_{k<j} \tilde{\phi}(t_{ij}, t_{ik})\, y_{ik}\right)^2 \tag{6}$$

The Smoothing Spline Model Space

$J(f)$ induces an orthogonal decomposition of $\mathcal{H}$:

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$$

$$\mathcal{H}_0 = \{f : \; J(f) = 0\} \text{ is the null space of } J,$$
$$\mathcal{H}_1 = \left\{f \in \mathcal{H} : ||f||^2 = J(f)\right\}$$

For the cubic smoothing spline defined on $\chi = [0, 1]$,

$$J(f) = \int_0^1 \left(f''(x)\right)^2 \; dx.$$

decomposes

$$\mathcal{H} = C^{(2)}[0, 1] = \left\{f : \int_0^1 \left(f''(x)\right)^2 \; dx < \infty\right\}$$

equipped with inner product

$$\langle f, g\rangle_{\mathcal{H}} = (M_0 f)(M_0 g) + (M_1 f)(M_1 g) + \int_0^1 f''(x)\, g''(x) \; dx,$$

ANOVA Decomposition of $\mathcal{H}$

$$\mathcal{H} = \underbrace{\mathcal{H}_{00}}_{\text{mean}} \oplus \underbrace{\mathcal{H}_{01}}_{\substack{\text{parametric}\\\text{constrast}}} \oplus \underbrace{\mathcal{H}_{1}}_{\substack{\text{nonparametric}\\\text{constrast}}}$$

where

$$\mathcal{H}_0 = \mathcal{H}_{00} \oplus \mathcal{H}_{01} = \{f : \ f \propto 1\} \oplus \{f : \ f \propto k_1\}$$

$$\mathcal{H}_1 = \{f : M_0 f = M_1 f = 0, \quad \int_0^1 \left(f''(x)\right)^2 \, dx < \infty\}.$$

The reproducing kernel $K = K_{00} + K_{01} + K_1$ can be defined in terms of the corresponding reproducing kernels

$$\begin{aligned} K_{00}(x,y) &= 1, \\ K_{01}(x,y) &= k_1(x)\,k_1(y)\,, \ \text{and} \\ K_1(x,y) &= k_2(x)\,k_2(y) - k_4(x-y)\,. \end{aligned}$$

where k_1, k_2, k_4 are the first, second, and fourth scaled Bernoulli polynomials.

The Tensor Product Smoothing Spline Space

Tensor Product Cubic Spline

	$\mathcal{H}_{00[2]}$	$\mathcal{H}_{01[2]}$	$\mathcal{H}_{1[2]}$
$\mathcal{H}_{00[1]}$	$\mathcal{H}_{00[1]} \otimes \mathcal{H}_{00[2]}$	$\mathcal{H}_{00[1]} \otimes \mathcal{H}_{01[2]}$	$\mathcal{H}_{00[1]} \otimes \mathcal{H}_{1[2]}$
$\mathcal{H}_{01[1]}$	$\mathcal{H}_{01[1]} \otimes \mathcal{H}_{00[2]}$	$\mathcal{H}_{01[1]} \otimes \mathcal{H}_{01[2]}$	$\mathcal{H}_{01[1]} \otimes \mathcal{H}_{1[2]}$
$\mathcal{H}_{1[1]}$	$\mathcal{H}_{1[1]} \otimes \mathcal{H}_{00[2]}$	$\mathcal{H}_{1[1]} \otimes \mathcal{H}_{01[2]}$	$\mathcal{H}_{1[1]} \otimes \mathcal{H}_{1[2]}$

	$\{1\}$	$\{k_1\}$	$\mathcal{H}_{1[2]}$
$\{1\}$	mean	p-main effect	np-main effect
$\{k_1\}$	p-main effect	$p \times p$-interaction	$p \times np$-interaction
$\mathcal{H}_{1[1]}$	np-main effect	$np \times p$-interaction	$np \times np$-interaction

A Reproducing Kernel Hilbert Space Framework for ϕ

Let

$$\begin{aligned}\phi \in \mathcal{H} &= \mathcal{H}_{[l]} \otimes \mathcal{H}_{[m]} \\ &= \mathcal{H}_0 \oplus \mathcal{H}_1.\end{aligned}$$

with RK $K = K_{[l]}K_{[m]}$. Fix $\sigma_{ij}^2 = \sigma^2\left(t_{ij}\right)$ in (6), find ϕ minimizing

$$-2\ell\left(\phi|Y_1,\ldots,Y_N,\sigma^2\right) + \lambda J\left(\phi\right) = \sum_{i=1}^{N}\sum_{j=2}^{p_i}\frac{1}{\sigma_{ij}^2}\left(y_{ij} - \sum_{k<j}\phi\left(\boldsymbol{v}_{ijk}\right)y_{ik}\right)^2 + \lambda||P_1\phi||^2 \tag{7}$$

where $J\left(\phi\right) = ||P_1\phi||^2$ and $\boldsymbol{v}_{ijk} \in \mathcal{V} = [0,1]^2$,

$$\begin{aligned}\boldsymbol{v}_{ijk} &= \left(t_{ij} - t_{ik}, \frac{1}{2}\left(t_{ij} + t_{ik}\right)\right) \\ &= \left(l_{ijk}, m_{ijk}\right)\end{aligned}$$

Define the set of unique within-subject pairs of observation times:

$$V = \bigcup_{i,j,k}\left\{\boldsymbol{v}_{ijk}\right\} \equiv \left\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{|V|}\right\}$$

A Representer Theorem

Theorem

Let $\{\nu_1, \ldots, \nu_{\mathcal{N}_0}\}$ *span* $\mathcal{H}_0$*, the null space of* $J(\phi) = ||P_1\phi||^2$*. Let* B *denote the* $|V| \times \mathcal{N}_0$ *matrix having* i^{th} *column equal to* ν_i *evaluated at the observed* $\boldsymbol{v} \in V$*, and assume that* B *has full column rank. Then the minimizer* ϕ_λ *of* (7) *is given by*

$$\phi_\lambda(\boldsymbol{v}) = \sum_{i=1}^{\mathcal{N}_0} d_i \nu_i(\boldsymbol{v}) + \sum_{j=1}^{|V|} c_j K_1(\boldsymbol{v}_j, \boldsymbol{v}), \tag{8}$$

where $K_1(\boldsymbol{v}_j, \boldsymbol{v})$ *denotes the reproducing kernel for* $\mathcal{H}_1$ *evaluated at* $\boldsymbol{v}_j$*, the* j^{th} *element of* V*.*

Obtaining the solution ϕ_λ

By the Representer Theorem 1, (7) becomes

$$-2\ell\left(c, d|\tilde{Y}, \tilde{B}, \tilde{K}_V\right)+\lambda J(\phi) = \left[\tilde{Y}-\tilde{B}d-\tilde{K}_V c\right]'\left[\tilde{Y}-\tilde{B}d-\tilde{K}_V c\right]+\lambda c' K_V c,$$

where

$$Y = (Y_1', Y_2', \ldots, Y_N')' = (y_{12}, y_{13}, \ldots, y_{1p_1}, \ldots, y_{N2}, \ldots, y_{Np_N})'$$
$$D = diag\left(\sigma_{12}^2, \sigma_{13}^2, \ldots, \sigma_{1p_1}^2, \ldots, \sigma_{N2}^2, \ldots, \sigma_{Np_N}^2\right)$$

$X_i = (p_i - 1) \times |V|$ matrix of AR covariates for Subject i

$K_V = |V| \times |V|$ matrix with (i, j) element $K_1(\boldsymbol{v}_i, \boldsymbol{v}_j)$

$B = |V| \times \mathcal{N}_0$ matrix with (i, j) element $\nu_j(\boldsymbol{v}_i)$

and $\tilde{Y} = D^{-1/2}Y$, $\tilde{B} = D^{-1/2}XB$, and $\tilde{K}_V = D^{-1/2}XK_V$,
$X = \begin{bmatrix} X_1' & X_2' & \ldots & X_N' \end{bmatrix}'$

Obtaining the solution ϕ_λ

Setting derivatives equal to zero, for fixed λ, c and d satisfy

$$\underbrace{\begin{bmatrix} \tilde{B}'\tilde{B} & \tilde{B}'\tilde{K}_V \\ \tilde{K}_V'\tilde{B} & \tilde{K}_V'\tilde{K}_V + \lambda K_V \end{bmatrix}}_{C'C} \begin{bmatrix} d \\ c \end{bmatrix} = \begin{bmatrix} \tilde{B}'\tilde{Y} \\ \tilde{K}_V'\tilde{Y} \end{bmatrix}$$

$$\implies C^{-1}(C')^{-1} \begin{bmatrix} \tilde{B}' \\ \tilde{K}_V' \end{bmatrix} \tilde{Y} = \begin{bmatrix} \hat{d} \\ \hat{c} \end{bmatrix}.$$

The fitted values are given by $\widehat{\tilde{Y}} = \tilde{A}_{\lambda,\boldsymbol{\theta}}\tilde{Y}$, where the smoothing matrix is

$$\tilde{A}_{\lambda,\boldsymbol{\theta}} = \begin{bmatrix} \tilde{B} & \tilde{K}_V \end{bmatrix} C^{-1}(C')^{-1} \begin{bmatrix} \tilde{B}' \\ \tilde{K}_V' \end{bmatrix}$$

Smoothing Parameter Selection

- Let $\mu_{ij} = E\left[y_{ij} | y_{i1}, \ldots, y_{i,j-1}\right] = \sum_{k<j} \phi\left(\boldsymbol{v}_{ijk}\right) y_{ik}$.

$$U\left(\lambda\right) = \tilde{Y}'\left(I - \tilde{A}_{\lambda,\boldsymbol{\theta}}\right)^2 \tilde{Y} + 2\,\mathrm{tr}\left(\tilde{A}_{\lambda,\boldsymbol{\theta}}\right)$$

  is an **unbiased estimator of the risk** $E\left[L\left(\lambda\right)\right] = E\left[\sum_{i=1}^{N} \sum_{j=1}^{p_i} \frac{1}{\sigma_{ij}^2}\left(\hat{y}_{ij} - \mu_{ij}\right)^2\right]$.

- Let $\widehat{\tilde{\mu}}_i^{[-i]}$ denote the estimate of $E\left[\tilde{Y}_i | X_i\right]$ based on the data when $\tilde{Y}_i$ is omitted. The **leave-one-subject-out cross validation score**

$$V_{loso}\left(\lambda\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\tilde{Y}_i - \widehat{\tilde{\mu}}_i^{[-i]}\right)'\left(\tilde{Y}_i - \widehat{\tilde{\mu}}_i^{[-i]}\right)$$

  approximates $MSPE = \frac{1}{N}\sum_{i=1}^{N} E\left[||\tilde{Y}_i^* - \widehat{\tilde{\mu}}_i||^2\right]$, where $\tilde{Y}_i^*$ denotes a vector of new observations $\tilde{y}_{i1}^*, \tilde{y}_{i1}^*, \ldots, \tilde{y}_{i,p_i}^*$.

A RKHS Framework for $\log \sigma^2$

Fixing ϕ in (6), let $Z_i = (z_{i1}, \ldots, z_{ip_i})'$, $z_{ij} = \epsilon_{ij}^2$, where

$$\epsilon_{ij} = y_{ij} - \sum_{k<j} \phi\left(\boldsymbol{v}_{ijk}\right) y_{ik}.$$

The log likelihood of the squared working innovations $Z_1, \ldots, Z_N$ coincides with a Gamma distribution with scale parameter $\alpha = 2$:

$$-2\ell\left(\sigma^2 | Z_1, \ldots, Z_N\right) = \sum_{i=1}^{N} \sum_{j=1}^{p_i} \eta_{ij} + \sum_{i=1}^{N} \sum_{j=1}^{p_i} z_{ij} e^{-\eta_{ij}},$$

where $\eta_{ij} = \eta\left(t_{ij}\right) = \log \sigma^2\left(t_{ij}\right)$.

A RKHS Framework for $\log \sigma^2$

Take the estimator of $\eta(t) = \log \sigma^2(t)$ to minimize

$$-2\ell(\eta|Z_1, \ldots, Z_N) + \lambda J(\eta) = \sum_{i=1}^{N} \sum_{j=1}^{p_i} \eta(t_{ij}) + \sum_{i=1}^{N} \sum_{j=1}^{p_i} z_{ij} e^{-\eta(t_{ij})} + \lambda J(\eta), \tag{9}$$

for $\eta \in \mathcal{H}$, where $J(\eta) = ||P_1 \eta||^2$. Let

$$\mathcal{T} = \bigcup_{i,j} \{t_{ij}\},$$

Theorem 1 gives that the minimizer of (9) has the form

$$\eta_\lambda(t) = \sum_{i=1}^{\mathcal{N}_0} d_i \nu_i(t) + \sum_{j=1}^{|\mathcal{T}|} c_j K_1(t_j, t), \tag{10}$$

where $\{\nu_i\}$ span $\mathcal{H}_0$, $K_1(t_j, t)$ is the RK for $\mathcal{H}_1$ evalutated at t_j, the j^{th} element of $\mathcal{T}$.

Simulation Studies

Simulation Conditions

I. Complete data

Σ	Model I, II, III, IV, V
N	$50, 100$
p	$10, 20, 30$

II. Unbalanced data, $N = 50$

Σ	Model I, II, III, IV, V
p	$10, 20$
% missing	$0, 0.1, 0.2, 0.3$

Performance is measured with loss functions

$$\Delta_1\left(\Sigma, \hat{\Sigma}\right) = tr\left(\left(\Sigma^{-1}\hat{\Sigma} - \mathrm{I}\right)^2\right), \quad \Delta_2\left(\Sigma, \hat{\Sigma}\right) = tr\left(\Sigma^{-1}\hat{\Sigma}\right) - log|\Sigma^{-1}\hat{\Sigma}| - p$$

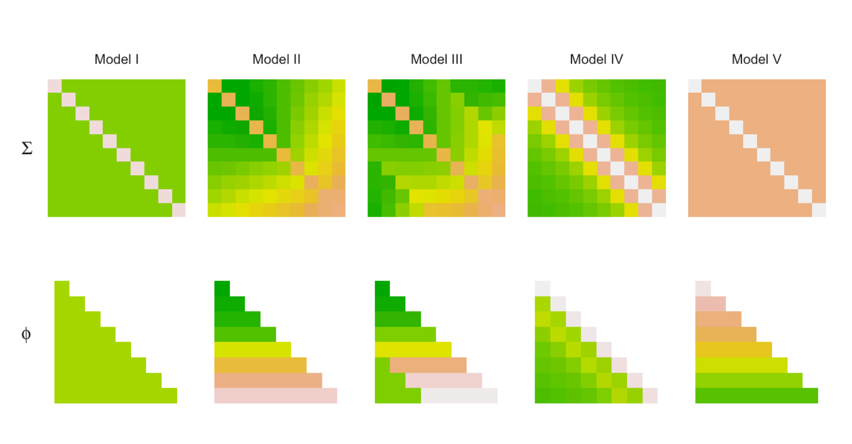
Use Monte Carlo simulation to estimate risk

$$R_i\left(\Sigma, \hat{\Sigma}\right) = E_\Sigma\left[\Delta_i\left(\Sigma, \hat{\Sigma}\right)\right], \quad i = 1, 2$$

In Study I, we compare to performance to that of the oracle estimator, Polynomial MCD GLM $\hat{\Sigma}_{poly}$, the sample covariance matrix $S = [s_{ij}]$, Shrinkage estimators S^ω, S^λ

Simulation Studies

Data Generation Models

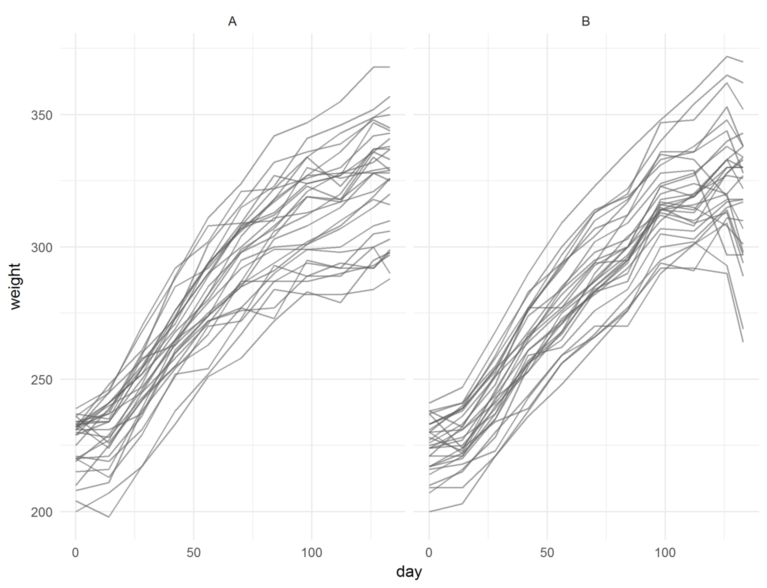


Simulation Studies

Data Generation Settings

I. $\Sigma = \mathrm{I}$	$\phi(t,s) = 0, 0 \le s < t \le 1$	$\sigma^2(t) = 1, 0 \le t \le 1$
II. $\Sigma = T^{-1} D T'^{-1}$	$\phi(t,s) = t - \frac{1}{2}, \quad 0 \le t \le 1$	$\sigma^2(t) = 0.1^2, 0 \le t \le 1$
III. $\Sigma = T^{-1} D T'^{-1}$	$\phi(t,s) = \begin{cases} t - \frac{1}{2}, & t - s \le 0.5 \\ 0, & t - s > 0.5 \end{cases}$	$\sigma^2(t) = 0.1^2, 0 \le t \le 1$
IV. $\Sigma = [\sigma_{ij}]$	$\sigma_{ij} = \left(1 + \frac{(t_i - t_j)^2}{2k^2}\right)^{-1}$	$k = 0.6, 0 < t_i, t_j < 1$
V. $\Sigma = \rho \mathrm{J} + (1-\rho)\mathrm{I}$, $\rho = 0.7$	$\phi_{ts} = \frac{\rho}{1+(t-2)\rho}, t = 2, \ldots, p$	$\sigma_t^2 = 1 - \frac{(\max(t,2)-2)\rho^2}{1+(\max(t,2)-2)\rho}$

Kenward Cattle Data



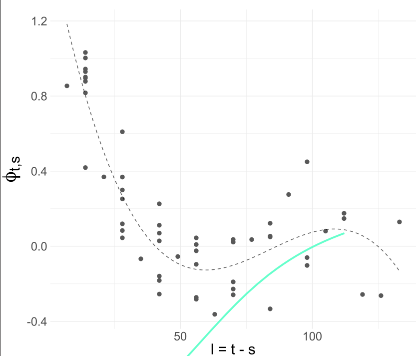
Kenward Cattle Data

Sample correlations

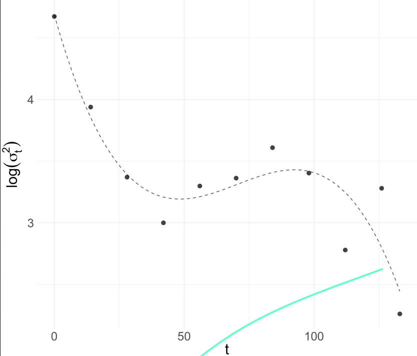
	day										
	0	14	28	42	56	70	84	98	112	126	133
0	1.00										
14	0.82	1.00									
28	0.76	0.91	1.00								
42	0.65	0.86	0.93	1.00							
56	0.63	0.83	0.89	0.93	1.00						
70	0.58	0.75	0.85	0.90	0.94	1.00					
84	0.51	0.64	0.75	0.80	0.85	0.92	1.00				
98	0.52	0.68	0.77	0.82	0.88	0.93	0.92	1.00			
112	0.51	0.61	0.71	0.74	0.81	0.89	0.92	0.96	1.00		
120	0.46	0.59	0.69	0.70	0.77	0.85	0.86	0.94	0.96	1.00	
133	0.46	0.56	0.67	0.67	0.74	0.81	0.84	0.91	0.95	0.98	1.00

The regressogram and innovation variogram

Sample generalized autoregressive parameters



Sample innovation variances



$$
\begin{aligned}
\phi_{ts} &= x_{ts}'\gamma \\
&= \gamma_0 + \gamma_1 (t-s) \\
&\qquad + \gamma_2 (t-s)^2 + \gamma_3 (t-s)^3
\end{aligned}
$$

$$
\begin{aligned}
\log \sigma_t^2 &= z_t'\xi \\
&= \left[\xi_0 + \xi_1 t + \xi_2 t^2 + \xi_3 t^3\right]
\end{aligned}
$$

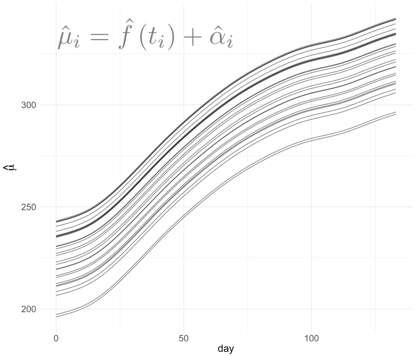
A joint mean-covariance model for the cattle weights

$$y_{ij} = f\left(t_i\right) + \alpha_i + \epsilon^*_{ij},$$
$$i = 1, \ldots, N = 30$$
$$j = 1, \ldots, 11,$$

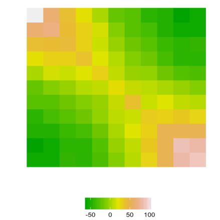
where the $\alpha_i \overset{\text{i.i.d.}}{\sim} N\left(0, \sigma^2_\alpha\right)$ mutually independent of $\epsilon^*_{ij} = \left(\epsilon^*_{i1}, \ldots, \epsilon^*_{ip_i}\right)'$,

$$\epsilon^*_i \sim N\left(0, \Sigma\right),$$
$$f \in \mathcal{H} = \mathcal{C}^2\left[\Re^+\right],$$

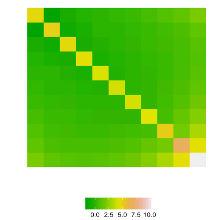
and $J\left(f\right) = \int_0^1 \left(f''\left(x\right)\right)^2 \, dx$.



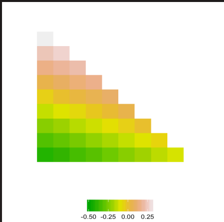
Modeling the Cholesky decomposition



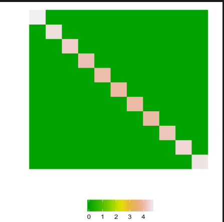
(a) S



(b) $\hat{\Sigma} = \hat{T'}^{-1} \hat{D} \hat{T}^{-1}$



(c) $\hat{\phi}\left(t, s\right)$



(d) $\hat{\sigma}^2\left(t\right)$

Model

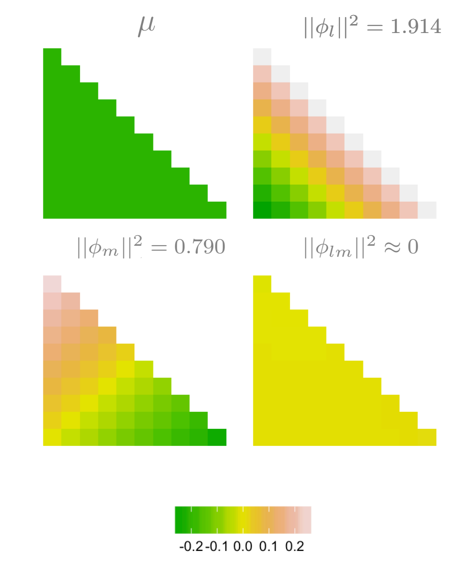
$$\epsilon^*\left(t_{ij}\right) = \sum_{k<j} \phi\left(t_{ij}, t_{ik}\right) \epsilon^*\left(t_{ij}\right) + \epsilon\left(t_{ij}\right)$$

where $\epsilon\left(t\right) \sim N\left(0, \sigma^2\left(t\right)\right)$ and

$$\phi \in \mathcal{H} = \mathcal{H}_{[l]} \otimes \mathcal{H}_{[m]}$$

$$\begin{aligned}
\mathcal{H}_{[l]} = & \left\{\phi : \ddot{\phi} = 0\right\} \\
& \oplus \left\{\phi : \phi\left(0\right), \dot{\phi}\left(0\right) = 0;\ \ddot{\phi} \in \mathcal{L}_2\left[0,1\right]\right\} \\
\mathcal{H}_{[m]} = & \left\{\phi : \phi \propto 1\right\} \\
& \oplus \left\{\phi : \int_0^1 \phi\, dx = 0,\ \ \dot{\phi} \in \mathcal{L}_2\left[0,1\right]\right\}
\end{aligned}$$

The functional components of ϕ



Concluding Remarks and Future Work

Thank you!

References I

Appendix

Tensor Product Cubic Spline Space

Reproducing Kernels

Subspace	Reproducing kernel
$\mathcal{H}_{00[1]} \otimes \mathcal{H}_{00[2]}$	1
$\mathcal{H}_{01[1]} \otimes \mathcal{H}_{00[2]}$	$k_1(x_1)\,k_1(y_1)$
$\mathcal{H}_{00[1]} \otimes \mathcal{H}_{01[2]}$	$k_1(x_2)\,k_1(y_2)$
$\mathcal{H}_{01[1]} \otimes \mathcal{H}_{01[2]}$	$k_1(x_1)\,k_1(y_1)\,k_1(x_2)\,k_1(y_2)$
$\mathcal{H}_{1[1]} \otimes \mathcal{H}_{00[2]}$	$k_2(x_1)\,k_2(y_1) - k_4(x_1 - y_1)$
$\mathcal{H}_{00[1]} \otimes \mathcal{H}_{1[2]}$	$k_2(x_2)\,k_2(y_2) - k_4(x_2 - y_2)$
$\mathcal{H}_{1[1]} \otimes \mathcal{H}_{01[2]}$	$[k_2(x_1)\,k_2(y_1) - k_4(x_1 - y_1)]\,k_1(x_2)\,k_1(y_2)$
$\mathcal{H}_{01[1]} \otimes \mathcal{H}_{1[2]}$	$k_1(x_1)\,k_1(y_1)\,[k_2(x_2)\,k_2(y_2) - k_4(x_2 - y_2)]$
$\mathcal{H}_{1[1]} \otimes \mathcal{H}_{1[2]}$	$[k_2(x_1)\,k_2(y_1) - k_4(x_1 - y_1)]\,[k_2(x_2)\,k_2(y_2) - k_4(x_2 - y_2)]$

A General Form for Multiple-Term RK Hilbert Spaces

Obtaining the solution ϕ_λ

When $\tilde{K}_V$ is not full rank

$$\begin{bmatrix} \tilde{B}'\tilde{B} & \tilde{B}'\tilde{K}_V \\ \tilde{K}_V'\tilde{B} & \tilde{K}_V'\tilde{K}_V + \lambda K_V \end{bmatrix} = \begin{bmatrix} C_1' & 0 \\ C_2' & C_3' \end{bmatrix} \begin{bmatrix} C_1 & C_2 \\ 0 & C_3 \end{bmatrix} \tag{11}$$

where $\tilde{B}'\tilde{B} = C_1'C_1$, $C_2 = (C_1')^{-1}\tilde{B}'\tilde{K}_V$, and $C_3'C_3 = \lambda K_V + \tilde{K}_V'\left(I - \tilde{B}\left(\tilde{B}'\tilde{B}\right)^{-1}\tilde{B}'\right)\tilde{K}_V$. Using an exchange of indices known as pivoting, one may write

$$C_3 = \begin{bmatrix} H_1 & H_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} H \\ 0 \end{bmatrix},$$

where H_1 is nonsingular. Define

$$\tilde{C}_3 = \begin{bmatrix} H_1 & H_2 \\ 0 & \delta I \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C_1 & C_2 \\ 0 & \tilde{C}_3 \end{bmatrix}; \tag{12}$$

Obtaining the solution ϕ_λ

When $\tilde{K}_V$ is not full rank

$$\begin{bmatrix} \tilde{B}'\tilde{B} & \tilde{B}'\tilde{K}_V \\ \tilde{K}_V'\tilde{B} & \tilde{K}_V'\tilde{K}_V + \lambda K_V \end{bmatrix} = \begin{bmatrix} C_1' & 0 \\ C_2' & C_3' \end{bmatrix} \begin{bmatrix} C_1 & C_2 \\ 0 & C_3 \end{bmatrix} \tag{13}$$

where $\tilde{B}'\tilde{B} = C_1'C_1$, $C_2 = (C_1')^{-1}\tilde{B}'\tilde{K}_V$, and $C_3'C_3 = \lambda K_V + \tilde{K}_V'\left(I - \tilde{B}\left(\tilde{B}'\tilde{B}\right)^{-1}\tilde{B}'\right)\tilde{K}_V$. Using an exchange of indices known as pivoting, one may write

$$C_3 = \begin{bmatrix} H_1 & H_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} H \\ 0 \end{bmatrix},$$

where H_1 is nonsingular. Define

$$\tilde{C}_3 = \begin{bmatrix} H_1 & H_2 \\ 0 & \delta I \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C_1 & C_2 \\ 0 & \tilde{C}_3 \end{bmatrix}; \tag{14}$$

Obtaining the solution ϕ_λ

When $\tilde{K}_V$ is not full rank

Then

$$\tilde{C}^{-1} = \begin{bmatrix} C_1^{-1} & -C_1^{-1} C_2 \tilde{C}_3^{-1} \\ 0 & \tilde{C}_3^{-1} \end{bmatrix}. \tag{15}$$

Premultiplying (13) by $(\tilde{C}')^{-1}$,

$$\begin{bmatrix} I & 0 \\ 0 & (\tilde{C}_3')^{-1} C_3' C_3 \tilde{C}_3^{-1} \end{bmatrix} \begin{bmatrix} \tilde{d} \\ \tilde{c} \end{bmatrix} = \begin{bmatrix} (C_1')^{-1} \tilde{B}' \tilde{Y} \\ (\tilde{C}_3')^{-1} \tilde{K}_V' \left(I - \tilde{B} \left(\tilde{B}' \tilde{B} \right)^{-1} \tilde{B}' \right) \tilde{Y} \end{bmatrix}$$

Minimizing U, V_{loso} with Multiple Smoothing Parameters

A RKHS Framework for $\log \sigma^2$

$$L(\eta) = \sum_{i=1}^{N} \sum_{j=1}^{p_i} \eta(t_{ij}) + \sum_{i=1}^{N} \sum_{j=1}^{p_i} z_{ij} e^{-\eta(t_{ij})} \qquad (16)$$

is continuous and convex in $\eta \in \mathcal{H}$. We assume that the $|V| \times \mathcal{N}_0$ matrix B which has (i, j) element $\nu_j(t_i)$ is full column rank, so that $L(\eta)$ is strictly convex in $\mathcal{H}$ and the minimizer of (9) uniquely exists.

Letting $\tilde{u}_{ij} = -z_{ij} e^{-\tilde{\eta}_{ij}}$, the Newton iteration uses the minimizer of the penalized weighted sums of squares

$$\sum_{i=1}^{N} \sum_{j=1}^{p_i} \left(\tilde{z}_{ij} - \eta(t_{ij})\right)^2 + \lambda J(\eta) \qquad (17)$$

to update $\tilde{\eta}$, where $\tilde{z}_{ij} = \tilde{\eta}(t_{ij}) - \tilde{u}_{ij}$.

Columns

Sometimes it's useful to split the screen

Test why is it gray?

Here's a column where I can write a bunch of things.

There are all sorts of things I can do in paragraph form.

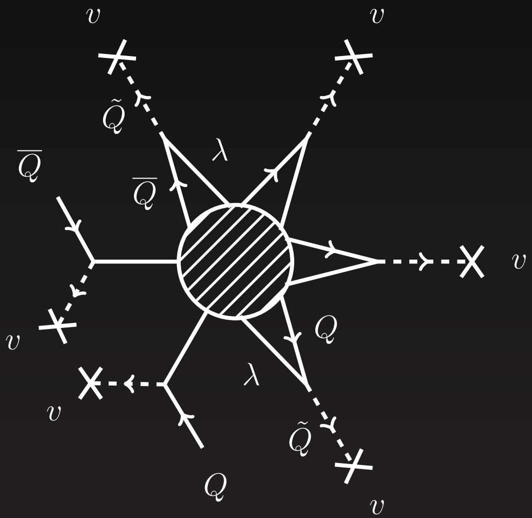
Blocks

...work in here too.

- Here's a column
- where I can itemize
- a bunch of things.

More Feynman diagrams

't Hooft operator



Drawing arrows onto a plot

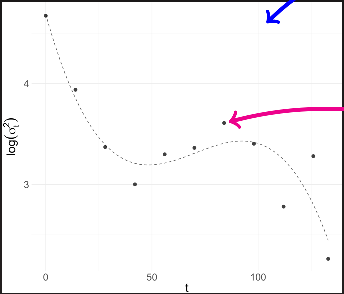
Well tempered neutralino

h^0 resonance

slepton co-annihilations

A^0 resonance

stop co-annihilations



Miscellaneous

- Use `\only<2>` to only show something for one overlay
- Can also use `<2->`
- For example, can highlight a word
- If you use `\uncover<3->` you get a   ... see?
- Protip: use `\textbackslash` to get a backslash

Miscellaneous

- Use `\only<2>` to only show something for one overlay
- Can also use `<2->`
- For example, can highlight highlight a word
- If you use `\uncover<3->` you get a ... see?
- Protip: use `\textbackslash` to get a backslash

Miscellaneous

- Use `\only<2>` to only show something for one overlay
- Can also use `<2->`
- For example, can highlight highlight a word
- If you use `\uncover<3->` you get a space ... see?
- Protip: use `\textbackslash` to get a backslash

Problems and Kludges

Things to work on

- There seems to be a bug in Beamer where the footnote color (defined using `setbeamercolor{footnote}` and `setbeamercolor{footnote mark}`) contaminates the normal text color. For now I suggest not using footnotes. They're of questionable use in a talk, anyway.
- Even though comment text is footnote-sized, it still has normal text line spacing. The `setspace` environment can fix this, but it forces a newline and it seems to make footnotes disappear.
- Make color theme more uniform and based on palette colors.

Problems and Kludges

XeLaTeX, LuaLaTeX

XeLaTeX doesn't allow one to use `setbeamertemplate[background canvas]` multiple times (e.g. to have one slide with a different background). A fix is to include `\def \pgfsysdriver{pgfsys-dvipdfmx.def}` before the `documentclass`, but this ends up breaking the arrows pointing to nodes.

In principle, LuaLaTeX can solve this, but that also requires some work since it only looks at Open Type Fonts (e.g. Gill Sans is not available by default).

`http://tex.stackexchange.com/questions/29497/xelatex-preventing-beamer-from-using-different-backgrounds`

Acknowledgements

I have borrowed heavily (and learned much) from Marco Barisione's Torino theme, which can be found one his blog. I have also learned and borrowed from Shawn Lankton's Keynote theme.

These can be found at

- `http://blog.barisione.org`
- `http://www.shawnlankton.com/2008/02/beamer-and-latex-with-keynote-theme/`

I've tried to maintain lots of comments in the `.tex` and `.sty` files to help other template-designers. At the moment it's all a jumbled mess, though!

Extra page: Additional hints

Look, it doesn't add to the total page count!

- Be sure to turn off any auto-notifiers (e.g. GMail)
- Consider using a PDF-to-keynote program; `http://www.cs.hmc.edu/~oneill/freesoftware/pdftokeynote.html`.
- Don't ever go over time.
- TikZ transparency trick: `http://www.texample.net/tikz/examples/transparent-png-overlay/`
- Use `addtocounter{framenumber}{-1}` for extra slides (like this one) to prevent it from screwing up the page numbering.