Supplementary material for Warped Functional Regression

Daniel Gervini
Department of Mathematical Sciences
University of Wisconsin–Milwaukee

September 3, 2014

Contents

Part I

Model

1 Identifiability

In this section we discuss the problem of identifiability of models of the form $x_i(t) = z_i\{w_i^{-1}(t)\}$ when the w_is belong to a given space of warping functions $\mathcal{W}$. We begin our discussion with a different but related issue: *when* is it possible to align curves in a meaningful way? Broadly speaking, curve alignment only makes sense when the curves have the same number of features (peaks and valleys). On the contrary, if for instance some curves in the sample have two peaks and some curves have three peaks, it is clearly not possible to align them in a meaningful way. In those cases it makes more sense to align the curves by clusters, as in Sangalli *et al.* (2010), but not to try to align the whole sample. So the first condition for a sensible curve alignment is that the curves have the same number of features. Once this is met, the existence and uniqueness of warping functions in $\mathcal{W}$ that perfectly align the curves is easy to establish. This discussion largely follows Kneip and Ramsay (2008), although we use a finite-dimensional $\mathcal{W}$.

Definition 1 *We say that the curves $\{x_1, \ldots, x_n\}$, with $x_i : [a, b] \to \mathbb{R}$ differentiable, have K features if*

$$\#\{t \in (a, b) : x_i'(t) = 0\} = K$$

for all $i = 1, \ldots, n$. In that case, for each curve $x_i(t)$ we define the vector of features $\boldsymbol{\zeta}_i = (\zeta_{i1}, \ldots, \zeta_{iK})$ as the zeros of $x_i'(t)$ taken in increasing order.

Definition 2 *We say that the curves* $\{z_1, \ldots, z_n\}$ *are aligned at* $\boldsymbol{\zeta}_0 \in \mathbb{R}^K$ *if they have* K *features and if* $\boldsymbol{\zeta}_i = \boldsymbol{\zeta}_0$ *for all* $i = 1, \ldots, n$.

Proposition 3 *If the curves* $\{x_1, \ldots, x_n\}$ *have* K *features and* $\mathcal{W}$ *is a* K*-dimensional linear space with basis* $\{\psi_1, \ldots, \psi_K\}$ *such that the vectors* $\{\psi_1(\bar{\boldsymbol{\zeta}}), \ldots, \psi_K(\bar{\boldsymbol{\zeta}})\}$ *are linearly independent, then there exist unique warping functions* $\{w_1, \ldots, w_n\}$ *in* $\mathcal{W}$ *such that* $\{x_1 \circ w_1, \ldots, x_n \circ w_n\}$ *are aligned at* $\bar{\boldsymbol{\zeta}}$.

Proof. The matrix $\boldsymbol{\Psi} = [\psi_1(\bar{\boldsymbol{\zeta}}), \ldots, \psi_K(\bar{\boldsymbol{\zeta}})] \in \mathbb{R}^{K \times K}$ is invertible by assumption. Let $w_i(t) = \sum_{j=1}^K \theta_{ij}\psi_j(t)$ with $\boldsymbol{\theta}_i = \boldsymbol{\Psi}^{-1}\boldsymbol{\zeta}_i$. Then the functions $\{x_1 \circ w_1, \ldots, x_n \circ w_n\}$ are aligned at $\bar{\boldsymbol{\zeta}}$, since

$$
\begin{aligned}
(x_i \circ w_i)'(\bar{\zeta}_{\cdot j}) &= x_i'(w_i(\bar{\zeta}_{\cdot j}))w_i'(\bar{\zeta}_{\cdot j}) \\
&= x_i'(\zeta_{ij})w_i'(\bar{\zeta}_{\cdot j}) \\
&= 0
\end{aligned}
$$

because the ζ_{ij}s are the zeros of $x_i'(t)$. Conversely, if $\{x_1 \circ \tilde{w}_1, \ldots, x_n \circ \tilde{w}_n\}$ are aligned at $\bar{\boldsymbol{\zeta}}$, then $(x_i \circ \tilde{w}_i)'(\bar{\zeta}_{\cdot j}) = 0$ for all i and j, which implies $x_i'(\tilde{w}_i(\bar{\zeta}_{\cdot j})) = 0$ for all i and j because $\tilde{w}_i'(\bar{\zeta}_{\cdot j}) > 0$. Then $\tilde{w}_i(\bar{\zeta}_{\cdot j}) = \zeta_{ij}$ for $j = 1, \ldots, K$, so if we write $\tilde{w}_i(t) = \sum_{j=1}^K \tilde{\theta}_{ij}\psi_j(t)$, we have $\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_i = \boldsymbol{\zeta}_i$. But $\boldsymbol{\Psi}$ is invertible, so $\tilde{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_i$ and then $\tilde{w}_i = w_i$ for all i. ∎

Note that $\{\psi_1, \ldots, \psi_K\}$ being linearly independent functions does not automatically imply that $\{\psi_1(\bar{\boldsymbol{\zeta}}), \ldots, \psi_K(\bar{\boldsymbol{\zeta}})\}$ are linearly independent vectors in $\mathbb{R}^K$. However, this will generally be the case if the family $\mathcal{W}$ is reasonably specified. Consider, for example, the case of monotone Hermite splines with knot vector $\boldsymbol{\tau}$. If $\boldsymbol{\tau} = \bar{\boldsymbol{\zeta}}$ then it is clear that $\{\psi_1(\bar{\boldsymbol{\zeta}}), \ldots, \psi_K(\bar{\boldsymbol{\zeta}})\}$ are independent, since $\psi_j(\boldsymbol{\tau}_0) = \tau_j\mathbf{e}_j$ by definition. But even if $\boldsymbol{\tau} \neq \bar{\boldsymbol{\zeta}}$, the vectors $\{\psi_1(\bar{\boldsymbol{\zeta}}), \ldots, \psi_K(\bar{\boldsymbol{\zeta}})\}$ will still be independent as long as $\boldsymbol{\tau}$ is reasonably close to $\bar{\boldsymbol{\zeta}}$. On the other hand, if $\boldsymbol{\tau}$ is grossly misspecified, this condition may fail: for example, if $\bar{\boldsymbol{\zeta}} = (.2, .3)$ and we take $\boldsymbol{\tau} = (.6, .7)$, given that the Hermite basis functions have compact support between three consecutive knots (or two knots and a boundary point), we have $\psi_2(\bar{\boldsymbol{\zeta}}) = \mathbf{0}$ and then $\{\psi_1(\bar{\boldsymbol{\zeta}}), \psi_2(\bar{\boldsymbol{\zeta}})\}$ are clearly not independent. But ruling out grossly misspecified cases like this, the condition that $\{\psi_1(\bar{\boldsymbol{\zeta}}), \ldots, \psi_K(\bar{\boldsymbol{\zeta}})\}$ are linearly independent is, in general, easily met.

The next proposition shows that the model $x_i(t) = z_i\{v_i(t)\}$ is identifiable under similar conditions to those of Proposition 3, although more general, since the dimension of $\mathcal{W}$ is allowed to be smaller than the number of features K (because, for example, the model may warp some peaks but not others, as in the simulations.)

Proposition 4 *Let* $\{x_1, \ldots, x_n\}$ *be a sample of curves generated by the model* $x_i(t) = z_i\{v_i(t)\}$*, where* $v_i(t) = w_i^{-1}(t)$*,* $w_i \in \mathcal{W}$*,* $\bar{w}(t) = t$ *for all* $t \in [a, b]$*, the curves* $\{z_1, \ldots, z_n\}$ *are aligned at* $\boldsymbol{\zeta}_0 \in \mathbb{R}^K$*, and* $\mathcal{W}$ *is an* r*-dimensional linear space with* $r \leq K$ *and basis* $\{\psi_1, \ldots, \psi_r\}$ *such that* $\{\psi_1(\boldsymbol{\zeta}_0), \ldots, \psi_r(\boldsymbol{\zeta}_0)\}$ *are linearly independent in* $\mathbb{R}^K$*. Then, if* $x_i(t) = \tilde{z}_i\{\tilde{v}_i(t)\}$ *for* $\tilde{v}_i(t) = \tilde{w}_i^{-1}(t)$ *with* $\tilde{w}_i \in \mathcal{W}$*,* $\overline{\tilde{w}}(t) = t$ *for all* $t \in [a, b]$*, and* $\{\tilde{z}_1, \ldots, \tilde{z}_n\}$ *aligned at some* $\tilde{\boldsymbol{\zeta}}_0 \in \mathbb{R}^{\tilde{K}}$*, we have* $\tilde{w}_i = w_i$ *and* $\tilde{z}_i = z_i$ *for all* i *(and in particular,* $\tilde{K} = K$ *and* $\tilde{\boldsymbol{\zeta}}_0 = \boldsymbol{\zeta}_0$*).*

Proof. By construction, $x_i(t)$ has K features given by $\boldsymbol{\zeta}_i = w_i(\boldsymbol{\zeta}_0)$, and $w_i(t) = \sum_{j=1}^r \theta_{ij}\psi_j(t)$ for some $\boldsymbol{\theta}_i \in \mathbb{R}^r$ that satisfies $\boldsymbol{\Psi}\boldsymbol{\theta}_i = \boldsymbol{\zeta}_i$, with $\boldsymbol{\Psi} = [\psi_1(\boldsymbol{\zeta}_0), \ldots, \psi_r(\boldsymbol{\zeta}_0)] \in \mathbb{R}^{K \times r}$ a full-rank matrix. Also, since $\bar{w}(t) \equiv t$, we have $\bar{\boldsymbol{\zeta}} = \boldsymbol{\zeta}_0$. Now suppose $x_i(t) = \tilde{z}_i\{\tilde{v}_i(t)\}$ with the $\tilde{z}_i(t)$s aligned at some $\tilde{\boldsymbol{\zeta}}_0$. Since $x_i'(t) = 0$ if and only if $\tilde{z}_i'\{\tilde{v}_i(t)\} = 0$, and we already know that $x_i'(t) = 0$ if and only if $t = \zeta_{ij}$ for some $j = 1, \ldots, K$, it follows that $\tilde{v}_i(\zeta_{ij})$ must be equal to one of the coordinates of $\tilde{\boldsymbol{\zeta}}_0$. But the $\tilde{v}_i$s are strictly increasing, so $\tilde{K}$ must be equal to K, and then $\tilde{v}_i(\boldsymbol{\zeta}_i) = \tilde{\boldsymbol{\zeta}}_0$, so $\tilde{w}_i(\tilde{\boldsymbol{\zeta}}_0) = \boldsymbol{\zeta}_i$ for all i. Taking averages we get $\tilde{\boldsymbol{\zeta}}_0 = \bar{\boldsymbol{\zeta}}$, since $\overline{\tilde{w}}(t) = t$ for all t; so $\tilde{\boldsymbol{\zeta}}_0 = \boldsymbol{\zeta}_0$. Then $\tilde{w}_i(\boldsymbol{\zeta}_0) = \boldsymbol{\zeta}_i$, that is, there exists a coefficient vector $\tilde{\boldsymbol{\theta}}_i \in \mathbb{R}^r$ such that $\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_i = \boldsymbol{\zeta}_i$; but $\boldsymbol{\Psi}$ is a full-rank matrix, so $\tilde{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_i$ and then $\tilde{w}_i = w_i$ for all i. Since $\tilde{z}_i(t) = x_i\{\tilde{w}_i(t)\}$ and $\tilde{w}_i = w_i$, it follows that $\tilde{z}_i = z_i$ for all i. ■

Part II

Estimation

1.1 Likelihood function

Under the distributional assumptions of § 2.2 in the paper, the likelihood function is derived as follows. The joint density function of the data vectors (x_i, y_i) and the latent random effects (w_i, z_i) can be factorized as

$$\begin{aligned} f(x_i, y_i, w_i, z_i) &= f(x_i, y_i \mid w_i, z_i) f(z_i \mid w_i) f(w_i) \\ &= f(x_i \mid w_i) f(y_i \mid z_i) f(z_i \mid w_i) f(w_i), \end{aligned}$$

since y_i depends on w_i only through z_i. Clearly $w_i \sim N(\mu_w, \Sigma_w)$ and $z_i | w_i \sim N\{\mu_z + A(w_i - \mu_w), \Sigma_e\}$. The conditional distributions of x_i given w_i and of y_i given z_i are derived as follows. Given $w_i = (u_i^T, \theta_{xi}^T)^T$ and $z_i = (v_i^T, \theta_{yi}^T)^T$, the values of θ_{xi} and θ_{yi}

determine the warping functions $\omega_i(s)$ and $\zeta_i(t)$ and consequently two warped time grids $s^*_{ij} = \omega_i^{-1}(s_{ij})$, $j = 1, \ldots, \nu_{1i}$, and $t^*_{ij} = \zeta_i^{-1}(t_{ij})$, $j = 1, \ldots, \nu_{2i}$. Let $B^*_{xi} \in \mathbb{R}^{\nu_{1i} \times q_1}$ and $B^*_{yi} \in \mathbb{R}^{\nu_{2i} \times q_2}$ be the B-spline bases evaluated at the warped time grids, that is $B^*_{xi,jk} = b_{xk}(s^*_{ij})$ and $B^*_{yi,jk} = b_{yk}(t^*_{ij})$. Then we have $x_i|w_i \sim N(B^*_{xi}m_x + B^*_{xi}Cu_i, \sigma^2_\varepsilon I_{\nu_{1i}})$ and $y_i|z_i \sim N(B^*_{yi}m_y + B^*_{yi}Dv_i, \sigma^2_\eta I_{\nu_{2i}})$. The maximum likelihood estimators maximize

$$\ell(A, \Sigma_e, \Sigma_w, m_x, m_y, C, D, \sigma^2_\varepsilon, \sigma^2_\eta) = \sum_{i=1}^{n} \log \iint f(x_i, y_i, w, z)\, \mathrm{d}w\, \mathrm{d}z$$

but the integrals do not have closed forms so we use the EM algorithm to find the optimum, treating the random effects (w_i, z_i) as missing data. This is explained next.

2 EM-algorithm

2.1 Log-likelihoods

$$-2 \log f(\mathbf{x}_i|\mathbf{w}_i) \propto \nu_{i1} \log \sigma^2_\varepsilon + \frac{\|\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}\mathbf{u}_i\|^2}{\sigma^2_\varepsilon},$$

$$-2 \log f(\mathbf{y}_i|\mathbf{z}_i) \propto \nu_{i2} \log \sigma^2_\eta + \frac{\|\mathbf{y}_i - \mathbf{B}_{yi}\mathbf{m}_y - \mathbf{B}_{yi}\mathbf{D}\mathbf{v}_i\|^2}{\sigma^2_\eta},$$

$$-2 \log f(\mathbf{z}_i|\mathbf{w}_i) \propto \sum_{j=1}^{p_2+r_2} \log \Sigma_{e,jj} + \sum_{j=1}^{p_2+r_2} \frac{\{\mathbf{z}_i - \boldsymbol{\mu}_z - \mathbf{A}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}^2_j}{\Sigma_{e,jj}},$$

$$-2 \log f(\mathbf{w}_i) \propto \log\det \boldsymbol{\Sigma}_w + (\mathbf{w}_i - \boldsymbol{\mu}_w)^T \boldsymbol{\Sigma}_w^{-1} (\mathbf{w}_i - \boldsymbol{\mu}_w).$$

2.2 E-step

Let $\mathbb{E}_i$ denote the conditional expectation of the random effects (missing data) given $(\mathbf{x}_i, \mathbf{y}_i)$ for the parameter $\hat{\boldsymbol{\omega}}^{(0)}$ of the current EM iteration (here $\boldsymbol{\omega}$ represents the collection of all model parameters). Then

$$Q(\boldsymbol{\omega}|\hat{\boldsymbol{\omega}}^{(0)}, \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n) := -2 \sum_{i=1}^{n} \mathbb{E}_i\{\log f(\mathbf{x}_i, \mathbf{y}_i, \mathbf{w}_i, \mathbf{z}_i)\}$$

satisfies

$$
\begin{aligned}
Q \quad \propto \quad & \left(\sum_{i=1}^{n}\nu_{i1}\right)\log\sigma_\varepsilon^2 + \frac{\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}\mathbf{u}_i\|^2\}}{\sigma_\varepsilon^2} \\
& + \left(\sum_{i=1}^{n}\nu_{i2}\right)\log\sigma_\eta^2 + \frac{\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{y}_i - \mathbf{B}_{yi}\mathbf{m}_y - \mathbf{B}_{yi}\mathbf{D}\mathbf{v}_i\|^2\}}{\sigma_\eta^2} \\
& + \sum_{j=1}^{p_2+r_2} n\log\Sigma_{e,jj} + \sum_{j=1}^{p_2+r_2}\frac{\sum_{i=1}^n \mathbb{E}_i[\{\mathbf{z}_i - \boldsymbol{\mu}_z - \mathbf{A}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}_j^2]}{\Sigma_{e,jj}} \\
& + n\log\det\boldsymbol{\Sigma}_w + \sum_{i=1}^{n}\mathbb{E}_i\{(\mathbf{w}_i - \boldsymbol{\mu}_w)^T\boldsymbol{\Sigma}_w^{-1}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}.
\end{aligned}
$$

2.3 M-step

The following updating equations are straightforward from above:

$$
\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{x}_i - \mathbf{B}_{xi}\hat{\mathbf{m}}_x - \mathbf{B}_{xi}\hat{\mathbf{C}}\mathbf{u}_i\|^2\}}{\sum_{i=1}^n \nu_{i1}},
$$

$$
\hat{\sigma}_\eta^2 = \frac{\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{y}_i - \mathbf{B}_{yi}\hat{\mathbf{m}}_y - \mathbf{B}_{yi}\hat{\mathbf{D}}\mathbf{v}_i\|^2\}}{\sum_{i=1}^n \nu_{i2}},
$$

$$
\hat{\Sigma}_{e,jj} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_i[\{\mathbf{z}_i - \hat{\boldsymbol{\mu}}_z - \hat{\mathbf{A}}(\mathbf{w}_i - \hat{\boldsymbol{\mu}}_w)\}_j^2],
$$

$$
\hat{\boldsymbol{\Sigma}}_w = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_i\{(\mathbf{w}_i - \hat{\boldsymbol{\mu}}_w)(\mathbf{w}_i - \hat{\boldsymbol{\mu}}_w)^T\}.
$$

Since $\hat{\boldsymbol{\Lambda}}$, the upper-left block of $\hat{\boldsymbol{\Sigma}}_w$, must be diagonal, we set

$$
(\hat{\boldsymbol{\Sigma}}_w)_{11} \longleftarrow \mathrm{diag}\{(\hat{\boldsymbol{\Sigma}}_w)_{11}\}.
$$

For $\mathbf{m}_x$ and $\mathbf{m}_y$ we have the estimating equations

$$
\begin{aligned}
\sum_{i=1}^{n}\mathbb{E}_i\{\mathbf{B}_{xi}^T(\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}\mathbf{u}_i)\} &= \mathbf{0}, \\
\sum_{i=1}^{n}\mathbb{E}_i\{\mathbf{B}_{yi}^T(\mathbf{y}_i - \mathbf{B}_{yi}\mathbf{m}_y - \mathbf{B}_{yi}\mathbf{D}\mathbf{v}_i)\} &= \mathbf{0},
\end{aligned}
$$

so

$$\hat{\mathbf{m}}_x = \left\{\sum_{i=1}^n \mathbb{E}_i(\mathbf{B}_{xi}^T\mathbf{B}_{xi})\right\}^{-1} \sum_{i=1}^n \mathbb{E}_i\{\mathbf{B}_{xi}^T(\mathbf{x}_i - \mathbf{B}_{xi}\hat{\mathbf{C}}\mathbf{u}_i)\},$$
$$\hat{\mathbf{m}}_y = \left\{\sum_{i=1}^n \mathbb{E}_i(\mathbf{B}_{yi}^T\mathbf{B}_{yi})\right\}^{-1} \sum_{i=1}^n \mathbb{E}_i\{\mathbf{B}_{yi}^T(\mathbf{y}_i - \mathbf{B}_{yi}\hat{\mathbf{D}}\mathbf{v}_i)\}.$$

For $\mathbf{C}$ and $\mathbf{D}$ we do the minimization sequentially because it's more complicated. Let $\mathbf{C}_{(k)}$ denote the matrix $\mathbf{C}$ without the kth column $\mathbf{c}_k$, and $\mathbf{u}_{i(k)}$ the vector $\mathbf{u}_i$ without coordinate u_{ik}. Then

$$\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}\mathbf{u}_i\|^2\} =$$

$$\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}_{(k)}\mathbf{u}_{i(k)}\|^2\}$$
$$+\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{B}_{xi}\mathbf{c}_k u_{ik}\|^2\}$$
$$-2\sum_{i=1}^n \mathbb{E}_i\{(\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}_{(k)}\mathbf{u}_{i(k)})^T\mathbf{B}_{xi}\mathbf{c}_k u_{ik}\}$$

with

$$\sum_{i=1}^n \mathbb{E}_i\{\|\mathbf{B}_{xi}\mathbf{c}_k u_{ik}\|^2\} = \mathbf{c}_k^T \left\{\sum_{i=1}^n \mathbb{E}_i(u_{ik}^2\mathbf{B}_{xi}^T\mathbf{B}_{xi})\right\}\mathbf{c}_k$$

and

$$\sum_{i=1}^n \mathbb{E}_i\{(\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}_{(k)}\mathbf{u}_{i(k)})^T\mathbf{B}_{xi}\mathbf{c}_k u_{ik}\} =$$
$$\left[\sum_{i=1}^n \mathbb{E}_i\{(\mathbf{x}_i - \mathbf{B}_{xi}\mathbf{m}_x - \mathbf{B}_{xi}\mathbf{C}_{(k)}\mathbf{u}_{i(k)})^T\mathbf{B}_{xi} u_{ik}\}\right]\mathbf{c}_k.$$

For each k, we have to minimize a quadratic function of the form $\mathbf{c}^T\boldsymbol{\Omega}_k\mathbf{c} - 2\mathbf{b}_k^T\mathbf{c}$ subject to the restrictions

$$\mathbf{c}^T\mathbf{J}_x\mathbf{c} = 1,$$
$$\mathbf{c}_j^T\mathbf{J}_x\mathbf{c} = 0 \text{ for } j < k.$$

Section 3 explains how this is done. Then we obtain $\hat{\mathbf{c}}_k$ for each $k = 1, \ldots, p_1$, and similarly with $\hat{\mathbf{d}}_k$ for $k = 1, \ldots, p_2$.

For $\mathbf{A}$, note that

$$\sum_{j=1}^{p_2+r_2} \frac{\sum_{i=1}^{n} \mathbb{E}_i[\{\mathbf{z}_i - \boldsymbol{\mu}_z - \mathbf{A}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}_j^2]}{\Sigma_{e,jj}}$$

$$= \sum_{i=1}^{n} \mathbb{E}_i[\|\boldsymbol{\Sigma}_e^{-1/2}\{\mathbf{z}_i - \boldsymbol{\mu}_z - \mathbf{A}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}\|^2].$$

Split the residual vector into the upper p_2-subvector and the lower r_2-subvector, and the diagonal matrix $\boldsymbol{\Sigma}_e$ accordingly. Then

$$\sum_{i=1}^{n} \mathbb{E}_i[\|\boldsymbol{\Sigma}_e^{-1/2}\{\mathbf{z}_i - \boldsymbol{\mu}_z - \mathbf{A}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}\|^2]$$

$$= \sum_{i=1}^{n} \mathbb{E}_i[\|\boldsymbol{\Sigma}_{e,11}^{-1/2}\{\mathbf{v}_i - \mathbf{A}_{1\cdot}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}\|^2]$$
$$+ \sum_{i=1}^{n} \mathbb{E}_i[\|\boldsymbol{\Sigma}_{e,22}^{-1/2}\{\boldsymbol{\theta}_{yi} - \boldsymbol{\mu}_{\theta_y} - \mathbf{A}_{2\cdot}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}\|^2].$$

The second term depends only on $\mathbf{A}_{2\cdot}$, which is not subject to restrictions, so it's minimized by

$$\hat{\mathbf{A}}_{2\cdot} = \sum_{i=1}^{n} \mathbb{E}_i\{(\boldsymbol{\theta}_{yi} - \boldsymbol{\mu}_{\theta_y})(\mathbf{w}_i - \boldsymbol{\mu}_w)^T\} \left[\sum_{i=1}^{n} \mathbb{E}_i\{(\mathbf{w}_i - \boldsymbol{\mu}_w)(\mathbf{w}_i - \boldsymbol{\mu}_w)^T\}\right]^{-1}$$

(the factor $\boldsymbol{\Sigma}_{e,22}^{-1/2}$ cancels out.)

The restriction that $\mathbf{A}_{1\cdot}\boldsymbol{\Sigma}_w\mathbf{A}_{1\cdot}^T$ be diagonal complicates minimization of the first term, which is carried out as explained in Section 4.

3 Constrained quadratic minimization

The basic problem is

$$\min \mathbf{c}^T\boldsymbol{\Omega}\mathbf{c} - 2\mathbf{b}^T\mathbf{c} \;\text{ s.t. }\; \|\mathbf{c}\|^2 = 1. \tag{1}$$

The Lagrangean is

$$L = \mathbf{c}^T\boldsymbol{\Omega}\mathbf{c} - 2\mathbf{b}^T\mathbf{c} + \ell(1 - \|\mathbf{c}\|^2),$$

so

$$\frac{\partial L}{\partial \mathbf{c}} = 2\boldsymbol{\Omega}\mathbf{c} - 2\mathbf{b} - 2\ell\mathbf{c} = \mathbf{0}$$

gives the solution

$$\mathbf{c}(\ell) = (\boldsymbol{\Omega} - \ell\mathbf{I})^{-1}\,\mathbf{b}$$

for ℓ such that $\|\mathbf{c}(\ell)\| = 1$.

To find ℓ: let $\boldsymbol{\Omega} = \mathbf{U}\mathbf{D}\mathbf{U}^T$ be the s.v.d. of $\boldsymbol{\Omega}$, and $\boldsymbol{\alpha} = \mathbf{U}^T\mathbf{b}$. Then

$$\hat{\mathbf{c}}(\ell) = \mathbf{U}(\mathbf{D} - \ell\mathbf{I})^{-1}\boldsymbol{\alpha}$$

and

$$\|\hat{\mathbf{c}}(\ell)\|^2 = \left\|(\mathbf{D} - \ell\mathbf{I})^{-1}\boldsymbol{\alpha}\right\|^2 = \sum_{j=1}^{q} \frac{\alpha_j^2}{(d_j - \ell)^2} =: g(\ell).$$

We have

$$\begin{aligned} g'(\ell) &= \sum_{j=1}^{q} \frac{2\alpha_j^2}{(d_j - \ell)^3}, \\ g''(\ell) &= \sum_{j=1}^{q} \frac{6\alpha_j^2}{(d_j - \ell)^4}, \end{aligned}$$

which are strictly positive in $(-\infty, \min(d_j))$, so $g(\ell)$ is strictly increasing and strictly convex in $(-\infty, \min(d_j))$. Since $g(\ell) \to 0$ when $\ell \to -\infty$ and $g(\ell) \to +\infty$ when $\ell \to \min(d_j)$, there exists a unique $\hat{\ell}$ such that $\|\hat{\mathbf{c}}(\hat{\ell})\|^2 = 1$. We find it iteratively using Newton–Raphson: given ℓ_0,

$$g(\ell_1) \approx g(\ell_0) + g'(\ell_0)(\ell_1 - \ell_0) = 1$$

if we set

$$\ell_1 = \ell_0 + \frac{(1 - g(\ell_0))}{g'(\ell_0)}.$$

Then iterate until convergence, using $\ell_0 = 0$ as initial guess.

The more general problem, with orthogonality constraints

$$\min \mathbf{c}^T\boldsymbol{\Omega}\mathbf{c} - 2\mathbf{b}^T\mathbf{c} \;\text{ s.t. }\; \|\mathbf{c}\|^2 = 1,\; \boldsymbol{\Gamma}\mathbf{c} = \mathbf{0} \tag{2}$$

where $\boldsymbol{\Gamma} \in \mathbb{R}^{m\times q}$ has full rank m ($m \leq q$) can be reduced to (1) as follows: let $\mathbf{V} \in \mathbb{R}^{q\times(q-m)}$ be an orthogonal matrix whose columns span the orthogonal complement of the rows of $\boldsymbol{\Gamma}$ (so $\boldsymbol{\Gamma}\mathbf{V} = \mathbf{O}$). Then $\boldsymbol{\Gamma}\mathbf{c} = \mathbf{0} \Leftrightarrow \mathbf{c} = \mathbf{V}\mathbf{a}$ for some $\mathbf{a} \in \mathbb{R}^{q-m}$. Since

$\|\mathbf{c}\|^2 = \|\mathbf{a}\|^2$, the solution to (2) is $\hat{\mathbf{c}} = \mathbf{V}\hat{\mathbf{a}}$ where

$$\hat{\mathbf{a}} = \arg\min \mathbf{a}^T\mathbf{V}^T\boldsymbol{\Omega}\mathbf{V}\mathbf{a} - 2\mathbf{b}^T\mathbf{V}\mathbf{a} \ \text{ s.t. } \ \|\mathbf{a}\|^2 = 1,$$

which is like (1).

4 Constrained LS problem

We want to

$$\min \sum_{i=1}^{n} \mathbb{E}_i[\|\boldsymbol{\Sigma}_{e,11}^{-1/2}\{\mathbf{v}_i - \mathbf{A}_{1\cdot}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}\|^2] \ \text{ s.t. } \mathbf{A}_{1\cdot}\boldsymbol{\Sigma}_w\mathbf{A}_{1\cdot}^T = \mathbf{D}$$

with $\mathbf{D} \in \mathbb{R}^{p_2 \times p_2}$ diagonal. Write

$$\begin{aligned}
& \sum_{i=1}^{n} \mathbb{E}_i[\|\boldsymbol{\Sigma}_{e,11}^{-1/2}\{\mathbf{v}_i - \mathbf{A}_{1\cdot}(\mathbf{w}_i - \boldsymbol{\mu}_w)\}\|^2] \\
= \ & \mathrm{tr}\left\{\sum_{i=1}^{n} \mathbb{E}_i\left(\boldsymbol{\Sigma}_{e,11}^{-1/2}\mathbf{v}_i\mathbf{v}_i^T\boldsymbol{\Sigma}_{e,11}^{-1/2}\right)\right\} \\
& + \mathrm{tr}\left[\sum_{i=1}^{n} \mathbb{E}_i\left\{\boldsymbol{\Sigma}_{e,11}^{-1/2}\mathbf{A}_{1\cdot}(\mathbf{w}_i - \boldsymbol{\mu}_w)(\mathbf{w}_i - \boldsymbol{\mu}_w)^T\mathbf{A}_{1\cdot}^T\boldsymbol{\Sigma}_{e,11}^{-1/2}\right\}\right] \\
& - 2\mathrm{tr}\left[\sum_{i=1}^{n} \mathbb{E}_i\left\{\boldsymbol{\Sigma}_{e,11}^{-1/2}\mathbf{v}_i(\mathbf{w}_i - \boldsymbol{\mu}_w)^T\mathbf{A}_{1\cdot}^T\boldsymbol{\Sigma}_{e,11}^{-1/2}\right\}\right] \\
& = n\mathrm{tr}\left\{\left(\mathbf{M}_1 - 2\mathbf{M}_2^T\mathbf{A}_{1\cdot}^T + \mathbf{A}_{1\cdot}\mathbf{M}_3\mathbf{A}_{1\cdot}^T\right)\boldsymbol{\Sigma}_{e,11}^{-1}\right\}
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{M}_1 &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_i\left(\mathbf{v}_i\mathbf{v}_i^T\right), \\
\mathbf{M}_2 &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_i\left\{(\mathbf{w}_i - \boldsymbol{\mu}_w)\mathbf{v}_i^T\right\}, \\
\mathbf{M}_3 &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_i\left\{(\mathbf{w}_i - \boldsymbol{\mu}_w)(\mathbf{w}_i - \boldsymbol{\mu}_w)^T\right\}.
\end{aligned}$$

But $\hat{\boldsymbol{\Sigma}}_w = \mathbf{M}_3$, so the problem comes down to

$$\min \mathrm{tr}\left\{\left(-2\mathbf{M}_2^T\mathbf{A}_{1\cdot}^T + \mathbf{A}_{1\cdot}\mathbf{M}_3\mathbf{A}_{1\cdot}^T\right)\boldsymbol{\Sigma}_{e,11}^{-1}\right\} \ \text{ s.t. } \mathbf{A}_{1\cdot}\mathbf{M}_3\mathbf{A}_{1\cdot}^T = \mathbf{D}. \tag{3}$$

Since $\mathbf{D}^{-1/2}\mathbf{A}_{1\cdot}\mathbf{M}_3\mathbf{A}_{1\cdot}^T\mathbf{D}^{-1/2} = \mathbf{I}_{p_2}$, the matrix $\mathbf{U} := \mathbf{M}_3^{1/2}\mathbf{A}_{1\cdot}^T\mathbf{D}^{-1/2} \in \mathbb{R}^{d_1 \times p_2}$ is orthogonal and $\mathbf{A}_{1\cdot}^T = \mathbf{M}_3^{-1/2}\mathbf{U}\mathbf{D}^{1/2}$. So (3) is equivalent to

$$\min \operatorname{tr}\left\{\left(-2\tilde{\mathbf{M}}_2^T\mathbf{U}\mathbf{D}^{1/2} + \mathbf{D}\right)\boldsymbol{\Sigma}_{e,11}^{-1}\right\} \text{ s.t. } \mathbf{U}^T\mathbf{U} = \mathbf{I}$$

where $\tilde{\mathbf{M}}_2 = \mathbf{M}_3^{-1/2}\mathbf{M}_2$.

Algorithm. Alternate between $\mathbf{D}$ and $\mathbf{U}$ until convergence. For fixed $\mathbf{D}$,

$$\begin{aligned}\hat{\mathbf{U}}(\mathbf{D}) &= \arg\max \operatorname{tr}\left(\tilde{\mathbf{M}}_2^T\mathbf{U}\mathbf{D}^{1/2}\boldsymbol{\Sigma}_{e,11}^{-1}\right) \\ &= \arg\max \operatorname{tr}\left(\mathbf{U}^T\tilde{\mathbf{M}}_2\boldsymbol{\Sigma}_{e,11}^{-1}\mathbf{D}^{1/2}\right)\end{aligned}$$

subject to $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. If $\tilde{\mathbf{U}}\tilde{\mathbf{S}}\tilde{\mathbf{U}}^T$ is the s.v.d. of $\tilde{\mathbf{M}}_2\boldsymbol{\Sigma}_{e,11}^{-1}\mathbf{D}\boldsymbol{\Sigma}_{e,11}^{-1}\tilde{\mathbf{M}}_2^T$, then $\hat{\mathbf{U}}(\mathbf{D})$ are the first p_2 columns of $\tilde{\mathbf{U}}$ (i.e. the ones corresponding to nonzero eigenvalues.)

For fixed $\mathbf{U}$,

$$\operatorname{tr}\left\{\left(-2\tilde{\mathbf{M}}_2^T\mathbf{U}\mathbf{D}^{1/2} + \mathbf{D}\right)\boldsymbol{\Sigma}_{e,11}^{-1}\right\} = \sum_{j=1}^{p_2}\frac{(-2\alpha_j\tilde{d}_j + \tilde{d}_j^2)}{\Sigma_{e,jj}}$$

where $\tilde{d}_j = d_j^{1/2}$ and $\alpha_j = (\tilde{\mathbf{M}}_2^T\mathbf{U})_{jj}$. This is minimized when $\tilde{d}_j = \alpha_j$, so

$$\hat{\mathbf{D}}^{1/2}(\mathbf{U}) = \operatorname{diag}(\tilde{\mathbf{M}}_2^T\mathbf{U}).$$

We iterate until convergence, starting with $\mathbf{D}_0^{1/2} = \boldsymbol{\Sigma}_{e,11}$ and $\mathbf{U}_0$ the s.v. factor of $\tilde{\mathbf{M}}_2\tilde{\mathbf{M}}_2^T$ (which is equivalent to starting with $\mathbf{A}_{1\cdot}^T = \mathbf{M}_3^{-1}\mathbf{M}_2$, the unconstrained minimizer.)

Part III

Asymptotics

5 Derivatives of log-likelihood function

The parameter is $\boldsymbol{\theta} = (\operatorname{vec}(\mathbf{A}^T)^T, \operatorname{v}(\boldsymbol{\Sigma}_w)^T)^T$; let $\boldsymbol{\theta}_1 = \operatorname{vec}(\mathbf{A}^T)$ and $\boldsymbol{\theta}_2 = \operatorname{v}(\boldsymbol{\Sigma}_w)$. We want to differentiate

$$f(\mathbf{x}, \mathbf{y}) = \int\int f(\mathbf{x}|\mathbf{w})f(\mathbf{y}|\mathbf{z})f(\mathbf{z}|\mathbf{w})f(\mathbf{w})\mathrm{d}\mathbf{w}\mathrm{d}\mathbf{z}$$

where only $f(\mathbf{z}|\mathbf{w})$ depends on $\mathbf{A}$ and only $f(\mathbf{w})$ depends on $\boldsymbol{\Sigma}_w$. To find the derivatives we use differentials, as in Magnus and Neudecker (1999).

Differentiating with respect to $\mathbf{A}$:

$$\mathrm{d}\{\log f(\mathbf{x},\mathbf{y})\} = \frac{\int\int f(\mathbf{x}|\mathbf{w})f(\mathbf{y}|\mathbf{z})\mathrm{d}\{f(\mathbf{z}|\mathbf{w})\}f(\mathbf{w})\mathrm{d}\mathbf{w}\mathrm{d}\mathbf{z}}{f(\mathbf{x},\mathbf{y})}.$$

Since

$$f(\mathbf{z}|\mathbf{w}) = const \times \exp\left\{-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu}_z-\mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w))^T\boldsymbol{\Sigma}_e^{-1}(\mathbf{z}-\boldsymbol{\mu}_z-\mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w))\right\}$$

we have

$$\begin{aligned}\mathrm{d}\{f(\mathbf{z}|\mathbf{w})\} &= const\times\exp\{\cdots\}\{(\mathbf{w}-\boldsymbol{\mu}_w)^T\mathrm{d}(\mathbf{A}^T)\boldsymbol{\Sigma}_e^{-1}(\mathbf{z}-\boldsymbol{\mu}_z-\mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w))\}\\ &= f(\mathbf{z}|\mathbf{w})\left\{(\mathbf{z}-\boldsymbol{\mu}_z-\mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w))^T\boldsymbol{\Sigma}_e^{-1}\otimes(\mathbf{w}-\boldsymbol{\mu}_w)^T\right\}\mathrm{vec}(\mathrm{d}(\mathbf{A}^T))\end{aligned}$$

and then

$$\nabla_{\boldsymbol{\theta}_1}f(\mathbf{z}|\mathbf{w}) = f(\mathbf{z}|\mathbf{w})\left\{\boldsymbol{\Sigma}_e^{-1}(\mathbf{z}-\boldsymbol{\mu}_z-\mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w))\otimes(\mathbf{w}-\boldsymbol{\mu}_w)\right\}.$$

Therefore

$$\nabla_{\boldsymbol{\theta}_1}\log f(\mathbf{x},\mathbf{y}) = \mathbb{E}\left\{\boldsymbol{\Sigma}_e^{-1}(\mathbf{z}-\boldsymbol{\mu}_z-\mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w))\otimes(\mathbf{w}-\boldsymbol{\mu}_w)\big|\,(\mathbf{x},\mathbf{y})\right\}.$$

Differentiating with respect to $\boldsymbol{\Sigma}_w$:

$$\mathrm{d}\{\log f(\mathbf{x},\mathbf{y})\} = \frac{\int\int f(\mathbf{x}|\mathbf{w})f(\mathbf{y}|\mathbf{z})f(\mathbf{z}|\mathbf{w})\mathrm{d}\{f(\mathbf{w})\}\mathrm{d}\mathbf{w}\mathrm{d}\mathbf{z}}{f(\mathbf{x},\mathbf{y})}.$$

Since

$$f(\mathbf{w}) = \frac{1}{(2\pi)^{d_1/2}\det(\boldsymbol{\Sigma}_w)^{1/2}}\exp\left\{-\frac{1}{2}(\mathbf{w}-\boldsymbol{\mu}_w)^T\boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w)\right\}$$

we have

$$\begin{aligned}\mathrm{d}\{f(\mathbf{w})\} &= \frac{1}{(2\pi)^{d_1/2}}\left[-\frac{1}{2}\det(\boldsymbol{\Sigma}_w)^{-3/2}\mathrm{d}\{\det(\boldsymbol{\Sigma}_w)\}\right]\exp\left\{-\frac{1}{2}(\mathbf{w}-\boldsymbol{\mu}_w)^T\boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w)\right\}\\ &\quad+\frac{1}{(2\pi)^{d_1/2}\det(\boldsymbol{\Sigma}_w)^{1/2}}\exp\left\{-\frac{1}{2}(\mathbf{w}-\boldsymbol{\mu}_w)^T\boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w)\right\}\\ &\quad\times\left\{-\frac{1}{2}(\mathbf{w}-\boldsymbol{\mu}_w)^T\mathrm{d}(\boldsymbol{\Sigma}_w^{-1})(\mathbf{w}-\boldsymbol{\mu}_w)\right\}\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{2} f(\mathbf{w}) \left[\det(\boldsymbol{\Sigma}_w)^{-1} \mathrm{d}\{\det(\boldsymbol{\Sigma}_w)\} + (\mathbf{w}-\boldsymbol{\mu}_w)^T \mathrm{d}(\boldsymbol{\Sigma}_w^{-1})(\mathbf{w}-\boldsymbol{\mu}_w)\right] \\
&= -\frac{1}{2} f(\mathbf{w}) \left[\mathrm{tr}\{\boldsymbol{\Sigma}_w^{-1}\mathrm{d}(\boldsymbol{\Sigma}_w)\} - (\mathbf{w}-\boldsymbol{\mu}_w)^T \boldsymbol{\Sigma}_w^{-1}\mathrm{d}(\boldsymbol{\Sigma}_w)\boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w)\right] \\
&= -\frac{1}{2} f(\mathbf{w}) \left[\mathrm{vec}(\boldsymbol{\Sigma}_w^{-1})^T - (\mathbf{w}-\boldsymbol{\mu}_w)^T \boldsymbol{\Sigma}_w^{-1} \otimes (\mathbf{w}-\boldsymbol{\mu}_w)^T \boldsymbol{\Sigma}_w^{-1}\right] \mathrm{vec}(\mathrm{d}(\boldsymbol{\Sigma}_w)).
\end{aligned}
$$

Since $\mathrm{vec}(\mathrm{d}(\boldsymbol{\Sigma}_w)) = \mathrm{d}(\mathrm{vec}(\boldsymbol{\Sigma}_w)) = \mathrm{d}(\mathbf{D}_{d_1}\mathrm{v}(\boldsymbol{\Sigma}_w))$, where $\mathbf{D}_{d_1}$ is the duplication matrix, we have

$$
\nabla_{\boldsymbol{\theta}_2} \log f(\mathbf{x},\mathbf{y}) = -\frac{1}{2}\mathbf{D}_{d_1}^T \mathbb{E}\left\{ \mathrm{vec}(\boldsymbol{\Sigma}_w^{-1}) - \boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w) \otimes \boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w) \middle| (\mathbf{x},\mathbf{y}) \right\}.
$$

Both $\nabla_{\boldsymbol{\theta}_1} \log f(\mathbf{x},\mathbf{y})$ and $\nabla_{\boldsymbol{\theta}_2} \log f(\mathbf{x},\mathbf{y})$ can be re-expressed in terms of $\mathbf{M}(\mathbf{x},\mathbf{y}) = \mathbb{E}\{(\mathbf{w}-\boldsymbol{\mu}_w)(\mathbf{w}-\boldsymbol{\mu}_w)^T|(\mathbf{x},\mathbf{y})\}$ and $\mathbf{N}(\mathbf{x},\mathbf{y}) = \mathbb{E}\{(\mathbf{w}-\boldsymbol{\mu}_w)(\mathbf{z}-\boldsymbol{\mu}_z)^T|(\mathbf{x},\mathbf{y})\}$ as follows:

$$
\begin{aligned}
\nabla_{\boldsymbol{\theta}_1} \log f(\mathbf{x},\mathbf{y}) &= \mathbb{E}\left\{ \boldsymbol{\Sigma}_e^{-1}(\mathbf{z}-\boldsymbol{\mu}_z - \mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w)) \otimes (\mathbf{w}-\boldsymbol{\mu}_w) \middle| (\mathbf{x},\mathbf{y}) \right\} \\
&= \mathrm{vec}\mathbb{E}\left\{ (\mathbf{w}-\boldsymbol{\mu}_w)(\mathbf{z}-\boldsymbol{\mu}_z - \mathbf{A}(\mathbf{w}-\boldsymbol{\mu}_w))^T \boldsymbol{\Sigma}_e^{-1} \middle| (\mathbf{x},\mathbf{y}) \right\} \\
&= \mathrm{vec}\left\{ \mathbf{N}(\mathbf{x},\mathbf{y})\boldsymbol{\Sigma}_e^{-1} - \mathbf{M}(\mathbf{x},\mathbf{y})\mathbf{A}^T\boldsymbol{\Sigma}_e^{-1} \right\}
\end{aligned}
$$

and

$$
\begin{aligned}
\nabla_{\boldsymbol{\theta}_2} \log f(\mathbf{x},\mathbf{y}) &= -\frac{1}{2}\mathbf{D}_{d_1}^T \mathbb{E}\left\{ \mathrm{vec}(\boldsymbol{\Sigma}_w^{-1}) - \boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w) \otimes \boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w) \middle| (\mathbf{x},\mathbf{y}) \right\} \\
&= -\frac{1}{2}\mathbf{D}_{d_1}^T \mathrm{vec}\mathbb{E}\left\{ \boldsymbol{\Sigma}_w^{-1} - \boldsymbol{\Sigma}_w^{-1}(\mathbf{w}-\boldsymbol{\mu}_w)(\mathbf{w}-\boldsymbol{\mu}_w)^T\boldsymbol{\Sigma}_w^{-1} \middle| (\mathbf{x},\mathbf{y}) \right\} \\
&= -\frac{1}{2}\mathbf{D}_{d_1}^T \mathrm{vec}\left\{ \boldsymbol{\Sigma}_w^{-1} - \boldsymbol{\Sigma}_w^{-1}\mathbf{M}(\mathbf{x},\mathbf{y})\boldsymbol{\Sigma}_w^{-1} \right\}.
\end{aligned}
$$

6 Derivatives of constraints

The condition that $\mathbf{A}_{1\cdot}\boldsymbol{\Sigma}_w\mathbf{A}_{1\cdot}^T$ be diagonal can be expressed as a system of $m = (p_2 - 1)p_2/2$ constraints of the form $h_{ij}(\boldsymbol{\theta}) = 0$, where $h_{ij}(\boldsymbol{\theta}) = \mathbf{a}_i^T\boldsymbol{\Sigma}_w\mathbf{a}_j$ and $\mathbf{a}_i^T$ is the ith row of $\mathbf{A}$, for $i = 2, \ldots, p_2$ and $j = 1, \ldots, i-1$. Now let $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \boldsymbol{\theta}_3^T)^T$ with $\boldsymbol{\theta}_1 = \mathrm{vec}(\mathbf{A}_{1\cdot}^T)$, $\boldsymbol{\theta}_2 = \mathrm{vec}(\mathbf{A}_{2\cdot}^T)$ and $\boldsymbol{\theta}_3 = \mathrm{v}(\boldsymbol{\Sigma}_w)$. Since $\mathbf{A}_{2\cdot}^T$ is not involved in any $h_{ij}(\boldsymbol{\theta})$, we have

$$
\nabla_{\boldsymbol{\theta}_2} h_{ij}(\boldsymbol{\theta}) = \mathbf{0}_{d_1 r_2}.
$$

Also, re-writing

$$\begin{aligned} h_{ij}(\boldsymbol{\theta}) &= (\mathbf{a}_j^T \otimes \mathbf{a}_i^T)\mathrm{vec}(\boldsymbol{\Sigma}_w) \\ &= (\mathbf{a}_j^T \otimes \mathbf{a}_i^T)\mathbf{D}_{d_1}\boldsymbol{\theta}_3 \end{aligned}$$

it's clear that

$$\nabla_{\boldsymbol{\theta}_3} h_{ij}(\boldsymbol{\theta}) = \mathbf{D}_{d_1}^T(\mathbf{a}_j \otimes \mathbf{a}_i).$$

Finally, since for any $i = 1, \ldots, p_2$ we have

$$\mathbf{a}_i = \mathbf{A}_{1\cdot}^T \mathbf{e}_{i,p_2} = (\mathbf{e}_{i,p_2}^T \otimes \mathbf{I}_{d_1})\mathrm{vec}(\mathbf{A}_{1\cdot}^T),$$

we can re-write

$$h_{ij}(\boldsymbol{\theta}) = \boldsymbol{\theta}_1^T(\mathbf{e}_{i,p_2} \otimes \mathbf{I}_{d_1})\boldsymbol{\Sigma}_w(\mathbf{e}_{j,p_2}^T \otimes \mathbf{I}_{d_1})\boldsymbol{\theta}_1$$

and then

$$\begin{aligned} h_{ij}(\boldsymbol{\theta}) &= \mathrm{d}(\boldsymbol{\theta}_1^T)(\mathbf{e}_{i,p_2} \otimes \mathbf{I}_{d_1})\boldsymbol{\Sigma}_w(\mathbf{e}_{j,p_2}^T \otimes \mathbf{I}_{d_1})\boldsymbol{\theta}_1 \\ &\quad +\boldsymbol{\theta}_1^T(\mathbf{e}_{i,p_2} \otimes \mathbf{I}_{d_1})\boldsymbol{\Sigma}_w(\mathbf{e}_{j,p_2}^T \otimes \mathbf{I}_{d_1})\mathrm{d}\boldsymbol{\theta}_1. \end{aligned}$$

Therefore

$$\begin{aligned} \nabla_{\boldsymbol{\theta}_1} h_{ij}(\boldsymbol{\theta}) &= (\mathbf{e}_{i,p_2} \otimes \mathbf{I}_{d_1})\boldsymbol{\Sigma}_w(\mathbf{e}_{j,p_2}^T \otimes \mathbf{I}_{d_1})\boldsymbol{\theta}_1 \\ &\quad +(\mathbf{e}_{j,p_2} \otimes \mathbf{I}_{d_1})\boldsymbol{\Sigma}_w(\mathbf{e}_{i,p_2}^T \otimes \mathbf{I}_{d_1})\boldsymbol{\theta}_1 \\ &= (\mathbf{e}_{i,p_2} \otimes \mathbf{I}_{d_1})\boldsymbol{\Sigma}_w\mathbf{a}_j + (\mathbf{e}_{j,p_2} \otimes \mathbf{I}_{d_1})\boldsymbol{\Sigma}_w\mathbf{a}_i \\ &= \mathrm{vec}\left(\boldsymbol{\Sigma}_w\mathbf{a}_j\mathbf{e}_{i,p_2}^T + \boldsymbol{\Sigma}_w\mathbf{a}_i\mathbf{e}_{j,p_2}^T\right) \\ &= \mathrm{vec}\left\{\boldsymbol{\Sigma}_w\mathbf{A}_{1\cdot}^T\left(\mathbf{e}_{j,p_2}\mathbf{e}_{i,p_2}^T + \mathbf{e}_{i,p_2}\mathbf{e}_{j,p_2}^T\right)\right\}. \end{aligned}$$

7 Proof of Theorem 1

This proof is a direct application of Theorem 4.4 of Geyer (1994); note that Theorem 5.2 of Geyer (1994), which pertains to consistent local minimizers instead of global minimizers, can also be applied because our $T_C(\xi_0)$ satisfies the stronger condition of being Clarke-regular (Rockafellar & Wets, 1998, ch. 6.B). Following Geyer's notation, let $F(\xi) = E\{-\log f(x, y;\xi)\}$ and $F_n(\xi) = -(1/n)\sum_{i=1}^n \log f(x_i, y_i;\xi)$. Then $\hat{\xi}_n = \arg\min_{\xi\in C} F_n(\xi)$

and $\xi_0 = \arg\min_{\xi \in C} F(\xi)$. Assumption A of Geyer (1994) is that

$$F(\xi) = F(\xi_0) + \frac{1}{2}(\xi - \xi_0)^T V(\xi - \xi_0) + o(\|\xi - \xi_0\|), \tag{4}$$

with $V = \nabla^2 F(\xi_0)$ positive definite. This is satisfied in our case because $\nabla F(\xi_0) = -E\{\nabla \log f(x, y;\xi_0)\} = 0$ and $\nabla^2 F(\xi_0) = E\{U(x, y)U(x, y)^T\}$. To see that the latter is positive definite, note that for $\xi = (\text{vec}(A^T)^T, \text{v}(\Sigma_w)^T)^T$ we have

$$\begin{aligned} U(x,y)^T\xi &= \text{tr}\{\Sigma_{e,0}^{-1}N(x,y)^T A^T\} - \text{tr}\{\Sigma_{e,0}^{-1}A_0 M(x,y)A^T\} \\ &\quad -\frac{1}{2}\text{tr}\{\Sigma_{w,0}^{-1}\Sigma_w - \Sigma_{w,0}^{-1}M(x,y)\Sigma_{w,0}^{-1}\Sigma_w\} \\ &= E\{(w-\mu_w)^T A^T \Sigma_{e,0}^{-1} e \mid (x,y)\} \\ &\quad -\frac{1}{2}\text{tr}(\Sigma_{w,0}^{-1}\Sigma_w) + \frac{1}{2}E\{(w-\mu_w)^T\Sigma_{w,0}^{-1}\Sigma_w\Sigma_{w,0}^{-1}(w-\mu_w) \mid (x,y)\}, \end{aligned}$$

where $e = z - \mu_z + A_0\,(w - \mu_w)$, then $\xi^T V\xi = E[\{U(x, y)^T\xi\}^2] \geq 0$ and it is equal to zero only if $U(x, y)^T\xi = 0$ with probability one, which can only happen if $\xi = 0$.

Assumption B of Geyer (1994), in our case, is that

$$-\log f(x, y; \xi) = -\log f(x, y; \xi_0) + (\xi - \xi_0)^T D(x, y) + \|\xi - \xi_0\| r(x, y, \xi)$$

for some $D(x, y)$ such that the remainder $r(x, y,\xi)$ is stochastically equicontinuous. This is satisfied by $D(x, y) = -\nabla \log f(x, y;\xi_0)$; the fact that $r(x, y,\xi)$ is stochastically equicontinuous follows from Pollard (1984, pp. 150–152). Clearly $D(x, y)$ satisfies a central limit theorem with asymptotic covariance matrix A that in this case is equal to V, so Assumption C of Geyer (1994) is also satisfied. Then Theorem 4.4 of Geyer (1994) can be applied. It states that the asymptotic distribution of $n^{1/2}(\hat{\xi}_n - \xi_0)$ is the same as the distribution of $\hat{\delta}(Z)$, the minimizer of

$$q_Z(\delta) = \delta^T Z + \frac{1}{2}\delta^T V\delta$$

over $\delta \in T_C(\xi_0)$, where $Z \sim N(0, A)$.

In our case $\hat{\delta}(Z)$ can be obtained in closed form, due to the simplicity of $T_C(\xi_0)$. Concretely, $T_C(\xi_0)$ is the space of δs such that $B\delta = 0$. Let $\Omega = (\Xi^*, \Xi)$ be a $d \times d$ orthogonal matrix whose first m columns Ξ^* span the space generated by the rows of B and whose last $d - m$ columns Ξ are orthogonal to the rows of B. Then $\delta \in T_C(\xi_0)$ if and only if $\delta = \Omega\beta$ with $\beta_1 = \cdots = \beta_m = 0$; that is, $\delta = \Xi\beta_2$ with β_2 the subvector containing the last $d-m$ coordinates of β. Then for $\delta \in T_C(\xi_0)$ we can write

$$q_Z(\delta) = \beta^T \Omega^T Z + \frac{1}{2}\beta^T \Omega^T V \Omega \beta = \beta_2^T \Xi^T Z + \frac{1}{2}\beta_2^T \Xi^T V \Xi \beta_2,$$

which is clearly minimized by $\hat{\beta}_2 = (\Xi^T V \Xi)^{-1} \Xi^T Z$. Therefore $\hat{\delta}(Z) = \Xi(\Xi^T V \Xi)^{-1}\Xi^T Z$, and since $A = V$, the result of the theorem follows.

Part IV

Simulations

8 Estimation and prediction accuracy

As explained in the paper, six models were simulated but only three reported. The full report is presented here in Tables 1 and 2, where "W" denotes warped functional regression and "O" ordinary functional regression. Also, we show some plots to give a better idea of the data that was generated. We show the underlying smooth trajectories $x_i(s)$ and $y_i(t)$ in Figures 1 and 2, and the actual data (i.e., the trajectories sampled at random time grids with added random noise) in Figures 3 and 4. Note that in Figures 3 and 4 the data are only the asterisks; the joining lines are just for better visualization. We also show the corresponding amplitude regression function $\beta(s,t)$ in Figure 5.

Note that models 1, 2 and 5 have the same amplitude regression function $\beta(s,t)$, and so do models 3, 4 and 6. The difference between models 1 and 2 is that, for the latter, warping and amplitude are correlated; the same goes for models 3 and 4. Models 1 and 2 have one-dimensional amplitude and warping processes (so the total dimension is 2), while models 3 and 4 have two-dimensional amplitude and warping processes (so the total dimension is 4). Models 5 and 6 follow non-Hermite warping processes, but the amplitude processes are like models 1 and 2, respectively.

9 Asymptotic accuracy study

In addition to the two simulation studies reported in the paper (estimation accuracy and prediction accuracy), we also studied the finite-sample adequacy of the asymptotic results of § 3 of the paper, particularly for hypothesis testing. We simulated data from Model 1 with $A = 0$, and also from a similar model that uses equally-spaced time grids of size 15

	Model 1				Model 2			
	bias		rmse		bias		rmse	
Parameter	W	O	W	O	W	O	W	O
β	0.12	0.19	0.21	0.30	0.11	0.69	0.33	0.74
μ_x	0.10	0.19	0.34	0.37	0.12	0.19	0.38	0.37
μ_y	0.13	0.32	0.42	0.51	0.16	0.59	0.49	0.73
ϕ_1	0.05	0.06	0.15	0.18	0.08	0.05	0.23	0.18
ψ_1	0.15	0.21	0.22	0.34	0.09	0.83	0.20	0.85
	Model 3				Model 4			
β	0.37	1.00	1.15	1.14	0.47	1.23	1.39	1.32
μ_x	0.14	0.27	0.46	0.47	0.13	0.26	0.47	0.46
μ_y	0.16	0.38	0.56	0.58	0.19	0.65	0.61	0.81
ϕ_1	0.92	0.99	1.23	1.40	0.96	0.99	1.36	1.40
ϕ_2	0.25	0.93	0.59	1.06	0.22	0.96	0.58	1.07
ψ_1	0.99	0.99	1.40	1.40	0.99	0.99	1.40	1.39
ψ_2	0.17	0.87	0.47	1.21	0.20	0.62	0.48	1.03
	Model 5				Model 6			
β	0.18	0.73	0.73	0.78	0.80	1.05	1.56	1.11
μ_x	0.44	0.94	0.84	1.10	0.55	0.94	0.93	1.11
μ_y	0.49	0.86	0.88	1.03	0.52	0.87	0.92	1.05
ϕ_1	0.18	0.68	0.50	0.86	0.98	0.99	1.39	1.40
ϕ_2	—	—	—	—	0.86	1.08	1.18	1.25
ψ_1	0.17	0.62	0.47	0.75	0.99	0.99	1.40	1.40
ψ_2	—	—	—	—	0.53	1.01	0.87	1.20

Table 1: Simulation Results. Bias and root mean squared errors for sample size $n = 50$

	Model 1				Model 2			
	bias		rmse		bias		rmse	
Parameter	W	O	W	O	W	O	W	O
β	0.12	0.18	0.18	0.24	0.12	0.70	0.29	0.72
μ_x	0.10	0.19	0.27	0.31	0.13	0.19	0.29	0.30
μ_y	0.14	0.33	0.32	0.43	0.19	0.60	0.40	0.68
ϕ_1	0.05	0.05	0.11	0.13	0.07	0.05	0.19	0.12
ψ_1	0.16	0.20	0.19	0.28	0.10	0.84	0.18	0.85
	Model 3				Model 4			
β	0.38	1.06	0.83	1.13	0.41	1.26	0.88	1.31
μ_x	0.13	0.27	0.34	0.38	0.11	0.27	0.34	0.38
μ_y	0.16	0.38	0.40	0.49	0.18	0.66	0.45	0.75
ϕ_1	0.55	0.99	0.79	1.40	0.48	0.99	0.70	1.40
ϕ_2	0.22	1.04	0.46	1.09	0.15	1.04	0.40	1.09
ψ_1	0.84	0.98	1.19	1.39	0.81	0.99	1.15	1.40
ψ_2	0.12	0.92	0.33	1.13	0.16	0.63	0.34	1.00
	Model 5				Model 6			
β	0.17	0.74	0.60	0.77	0.85	1.05	1.25	1.08
μ_x	0.43	0.95	0.69	1.04	0.53	0.95	0.74	1.03
μ_y	0.48	0.88	0.73	0.97	0.50	0.88	0.74	0.97
ϕ_1	0.15	0.76	0.42	0.87	0.99	0.99	1.40	1.40
ϕ_2	—	—	—	—	0.92	1.18	1.13	1.27
ψ_1	0.16	0.66	0.40	0.72	0.97	0.99	1.38	1.40
ψ_2	—	—	—	—	0.47	1.14	0.70	1.23

Table 2: Simulation Results. Bias and root mean squared errors for sample size $n = 100$

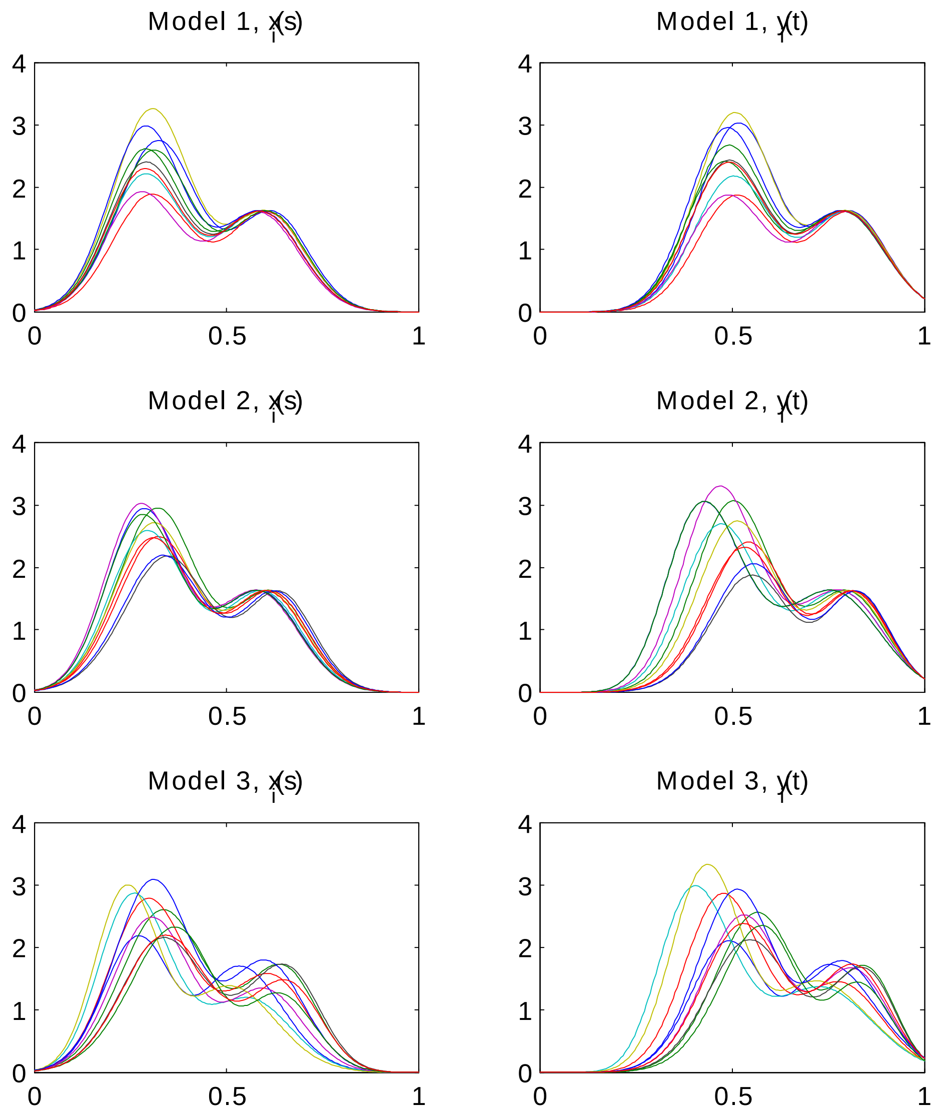

Figure 1: Some simulated covariate and response curves, models 1–3; these are the underlying (unobservable) smooth curves.

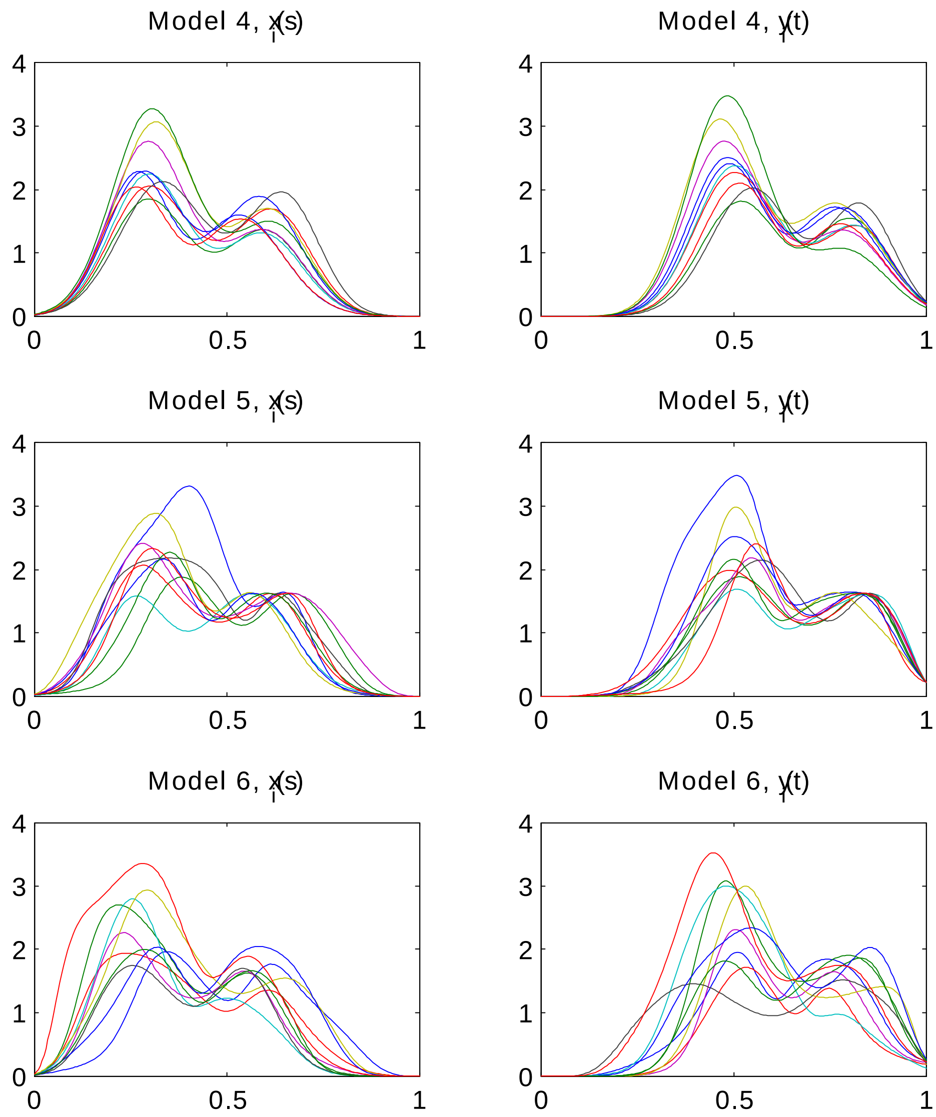

Figure 2: Some simulated covariate and response curves, models 4–6; these are the underlying (unobservable) smooth curves.

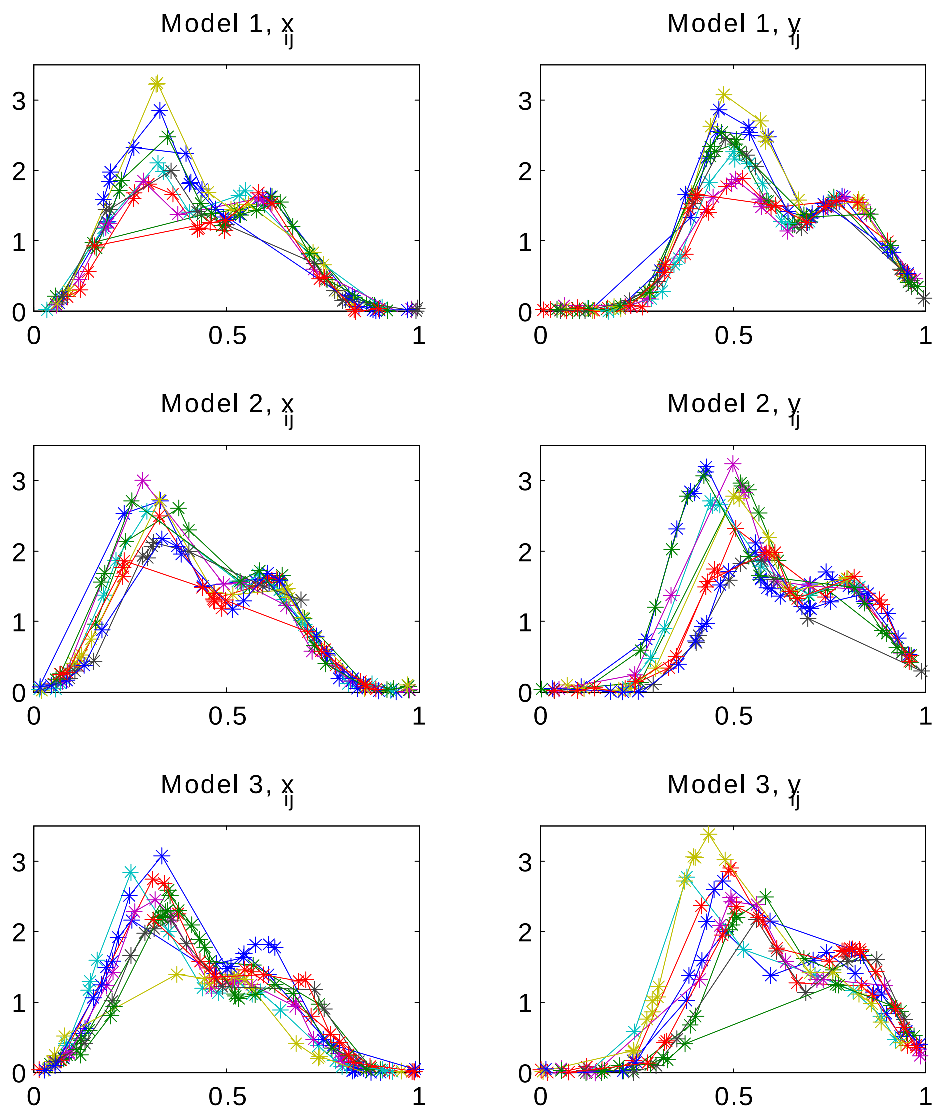

Figure 3: Some simulated covariate and response curves, models 1–3; these are the raw observations (the asterisks).

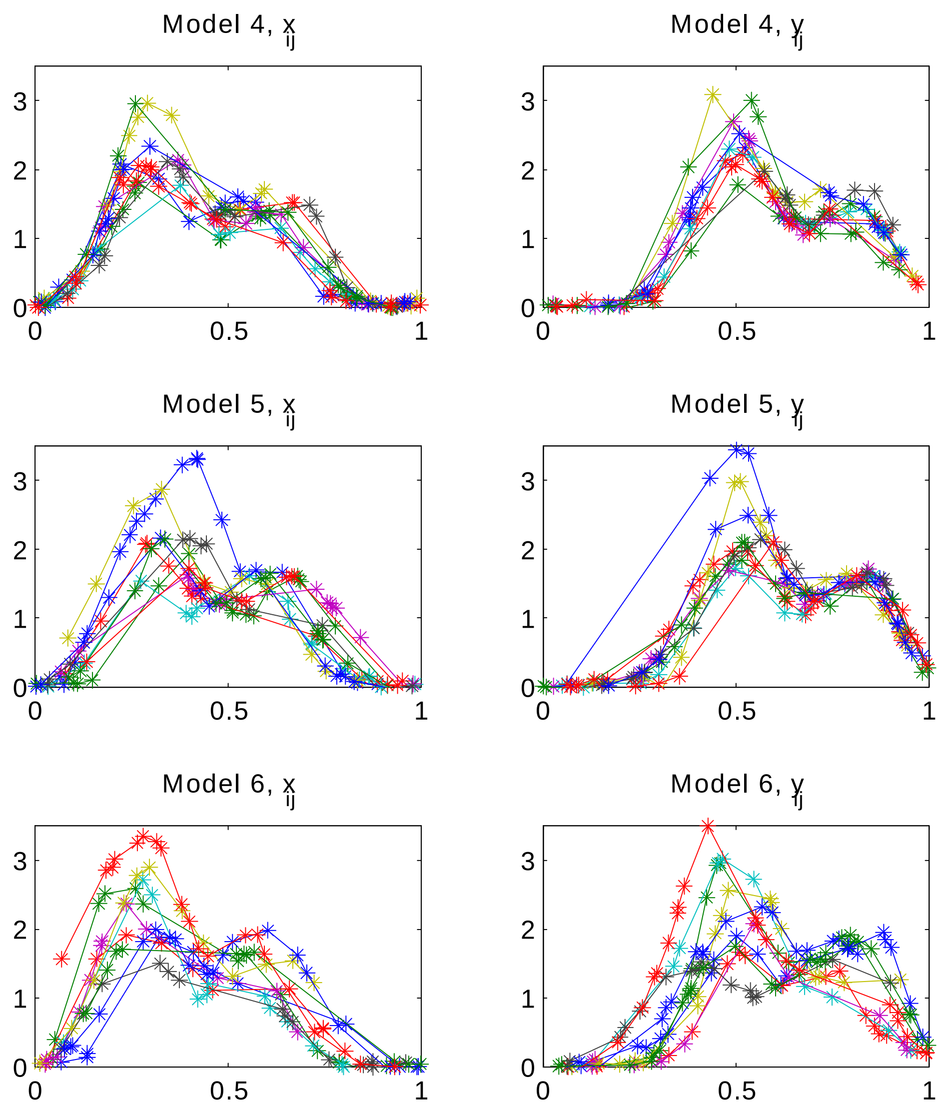

Figure 4: Some simulated covariate and response curves, models 4–6; these are the raw observations (the asterisks).

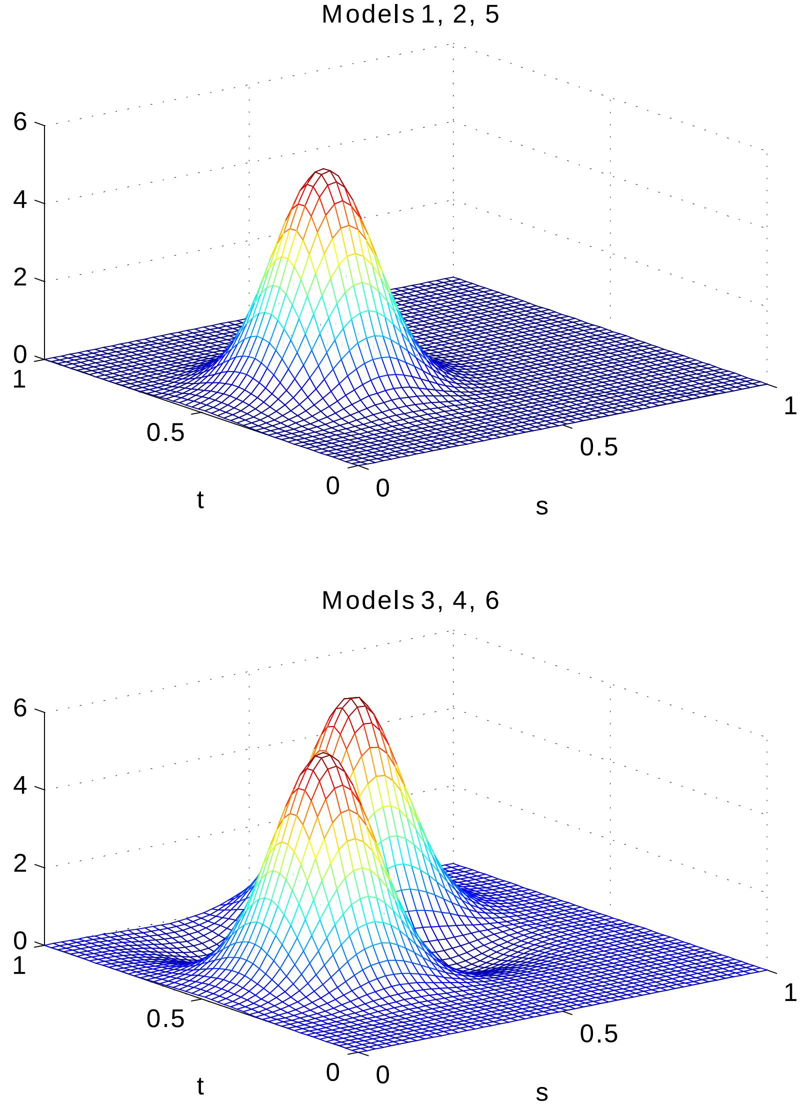

Figure 5: Amplitude regression function $\beta(s,t)$ used for the simulations.

	Random grids					
	$n = 50$			$n = 200$		
	Q	Z_{11}	Z_{12}	Q	Z_{11}	Z_{12}
True variance	0.09	0.08	0.08	0.10	0.10	0.09
Asymptotic	0.34	0.24	0.21	0.33	0.21	0.20
Bootstrap	0.25	0.16	0.13	0.25	0.14	0.11
	Equally spaced grids					
True variance	0.11	0.08	0.08	0.10	0.10	0.07
Asymptotic	0.36	0.20	0.23	0.27	0.14	0.26
Bootstrap	0.33	0.18	0.21	0.29	0.11	0.23

Table 3: Simulation Results. Tail probabilities of nominal 10%-level tests

instead of the random time grids of Model 1. Two sample sizes were considered in each case, $n = 50$ and $n = 200$. Each scenario was replicated 500 times.

The warped regression estimator was computed using the same specifications as above. The covariance matrix of $\mathrm{vec}(\hat{A}^T)$ was estimated by the asymptotic formulas of § 3 of the paper and by a bootstrap, using 50 bootstrap samples. The true covariance matrix of $\mathrm{vec}(\hat{A}^T)$ was computed as the sample covariance of the 500 replicated estimators. Since we are interested in testing, we computed tail probabilities of $Q = \mathrm{vec}(\hat{A}^T)^T \hat{\Sigma}^{-1} \mathrm{vec}(\hat{A}^T)$, where $\hat{\Sigma}$ is the respective covariance estimator of $\mathrm{vec}(\hat{A}^T)$, and of $Z_{1j} = \hat{a}_{1j} / \widehat{\mathrm{sd}}(\hat{a}_{1j})$ for $j = 1, 2$. Specifically, we report $P\,(Q \geq 7.78)$ and $P\,(|Z_{1j}| \geq 1.645)$ for $j = 1, 2$, which should be close to 0.10.

Table 3 shows the results. There are two aspects of the asymptotics that we are trying to assess: the adequacy of the normal approximation and the adequacy of the variance estimators. The first aspect can be best assessed using the true variance in the test statistics, so the variance estimation error is not a confounding factor. In this regard we see in Table 3 that the asymptotic approximation is good even for $n = 50$, both for the global Q-test and for the marginal Z-tests. In the more realistic cases where the variance is estimated, we see that bootstrap variance estimators generally work better than the asymptotic-variance formula; although both underestimate the true variances, bootstrap tends to underestimate them less, especially for random time grids.

Part V

Data Analysis

10 Zurich Longitudinal Growth data

Human height is the sum of leg length plus sitting height (head, neck and torso). The ratio of sitting height to leg length is about 1.1 for a healthy adult male. During adolescence, however, legs experience the growth spurt somewhat earlier than the trunk and the ratio temporarily falls below the ultimate adult value. Moreover, leg growth tappers off faster than trunk growth; trunk growth normally continues well after the 20th year of age for males (see Figure 6). The common impression that many teenagers have unusually long legs is then supported by the data, and generally does not indicate any growth pathology. However, a extreme sitting height to leg length ratio *might* indicate a growth pathology, so it is useful to investigate the relationship between leg and sitting height growth. Here we do this by fitting functional regression models using leg growth velocity as covariate and sitting heigth growth velocity as response. The data we use is the First Zurich Longitudinal Study of human growth (Prader et al., 1989), and we focus on the subset of 120 males.

We fitted two types of models: ordinary functional linear regression and warped functional regression, with varying numbers of principal components. As spline basis for the functional parameters we used, for all models, cubic B-splines with 5 equally spaced knots in the interval $[10, 22]$. For the warped regression models we used $\tau_{0x} = 13.5$ and $\tau_{0y} = 14.5$ as warping knots for the covariates and the responses, respectively, which are roughly the average peak locations in Figures 6(a) and 6(b). We considered three ordinary linear models and two warped regression models with increasing numbers of principal components (p_1, p_2), and computed crossvalidation prediction errors to find the optimal combination for each family. For ordinary linear regression we obtained the following crossvalidation errors: 3.67 for $(p_1, p_2) = (1, 1)$, 3.03 for $(2, 2)$ and 3.60 for $(3, 3)$, the optimal then being the four-dimensional model with two components for each of covariates and responses. For warped linear regression the crossvalidation errors were 2.74 for $(p_1, p_2) = (1, 1)$ and 2.85 for $(2, 2)$, the optimal being the first one, which is also four-dimensional (since in addition to the two amplitude components it has one-dimensional warping components for each of covariates and responses). Note that the crossvalidation error of the optimal warped model is 10% smaller than that of the ordinary regression model.

Figure 7 shows 20 response curves and the corresponding predicted curves, to give a

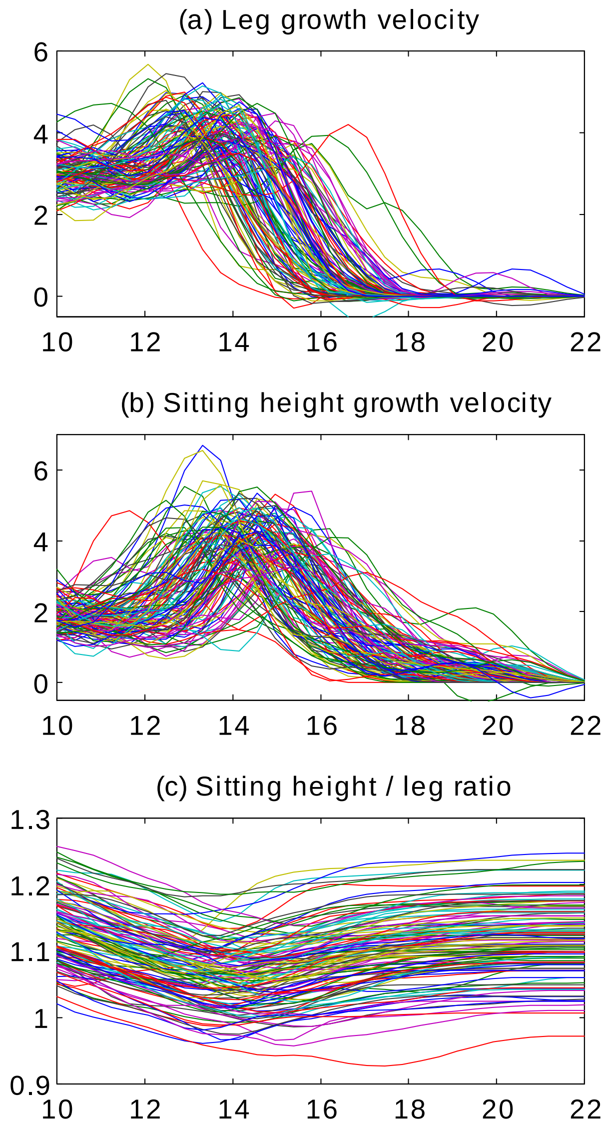

Figure 6: Zurich Growth Data Example. Leg growth velocity, sitting height growth velocity, and leg length to sitting height ratio for males.

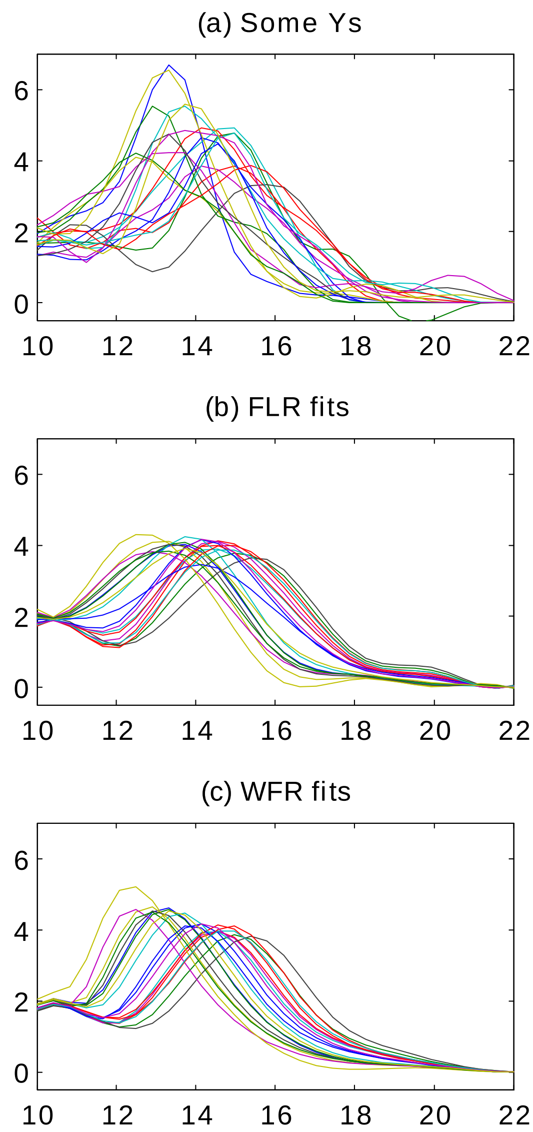

Figure 7: Zurich Growth Data Example. A subsample of response curves and corresponding predicted responses.

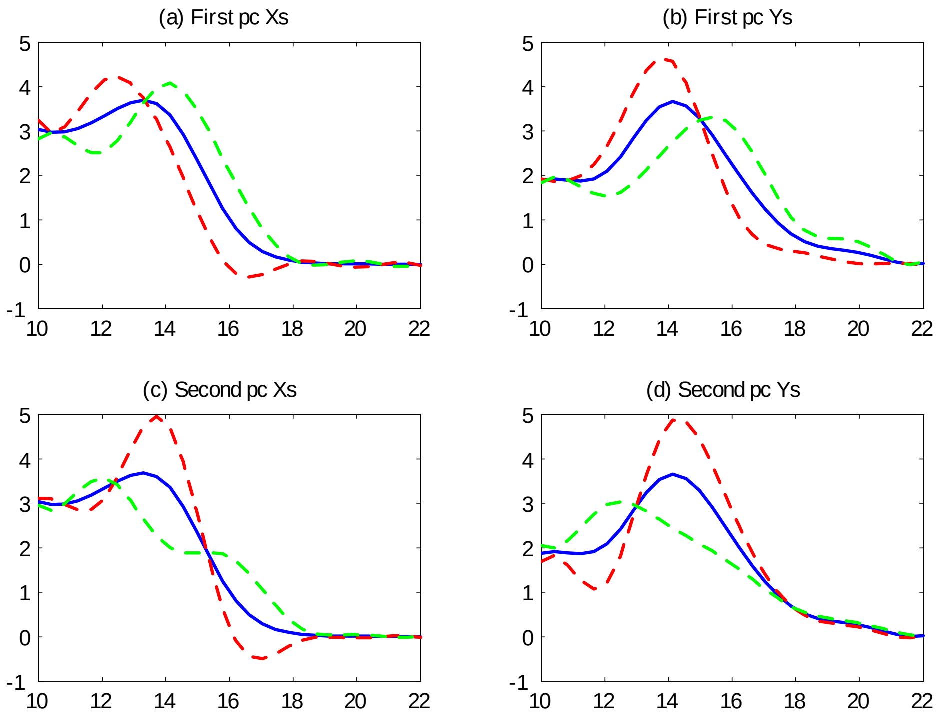

Figure 8: Zurich Growth Data Example. Principal components obtained by ordinary functional regression.

general idea of the kind of fits the methods produce. Qualitatively speaking, warped regression produces predictors with sharper peaks that more closely resemble the observed curves, whereas ordinary regression produces wider, blunter peaks, a result of the "smearing" effect of phase variability.

Regarding the principal component estimators, we see in Figure 8 that the first pc's produced by ordinary functional regression are mostly phase-variability related, almost exclusively so for the covariates (Figure 8(a)) and confounded with amplitude variability for the ys (Figure 8(b)). Warped functional regression, in contrast, produces two neat amplitude-variability components (Figure 9).

Warped regression also provides simple answers to questions like 'Can the timing of the sitting-height peak be predicted from the timing of the leg-growth peak?'. By looking at a scatter plot of the $\hat{\tau}_{yi}$s versus the $\hat{\tau}_{xi}$s (Figure 10(d)) we see that the answer is 'Yes, and strongly so'. Answering this question with ordinary functional regression is more complicated: one would look at the scatter plot of the first-pc component scores of the ys versus the first-pc component scores of the xs and see if there is a strong association,

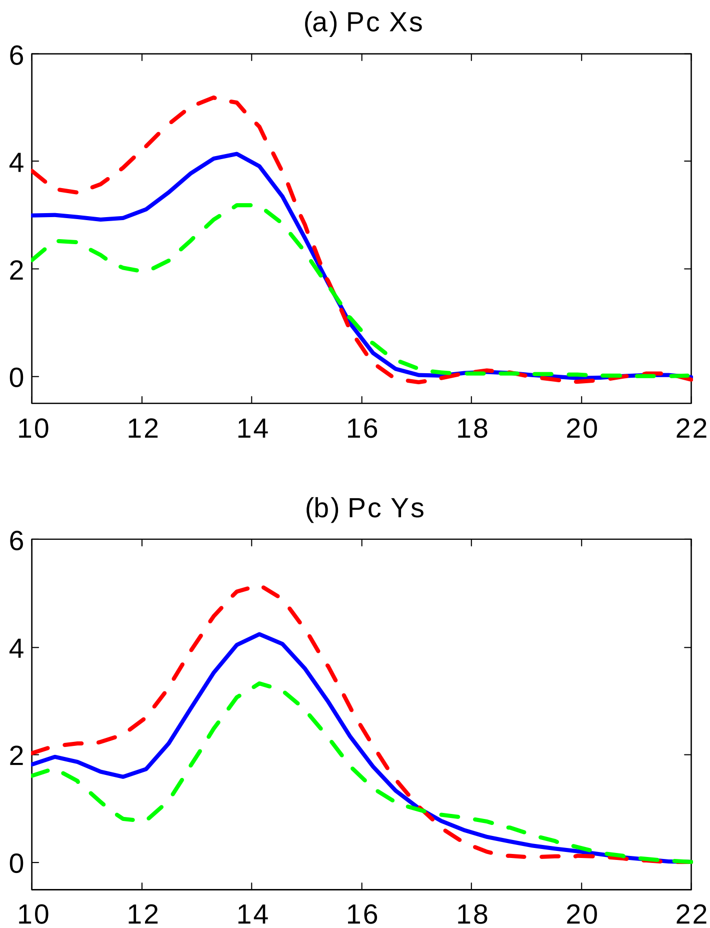

Figure 9: Zurich Growth Data Example. Principal components obtained by warped functional regression.

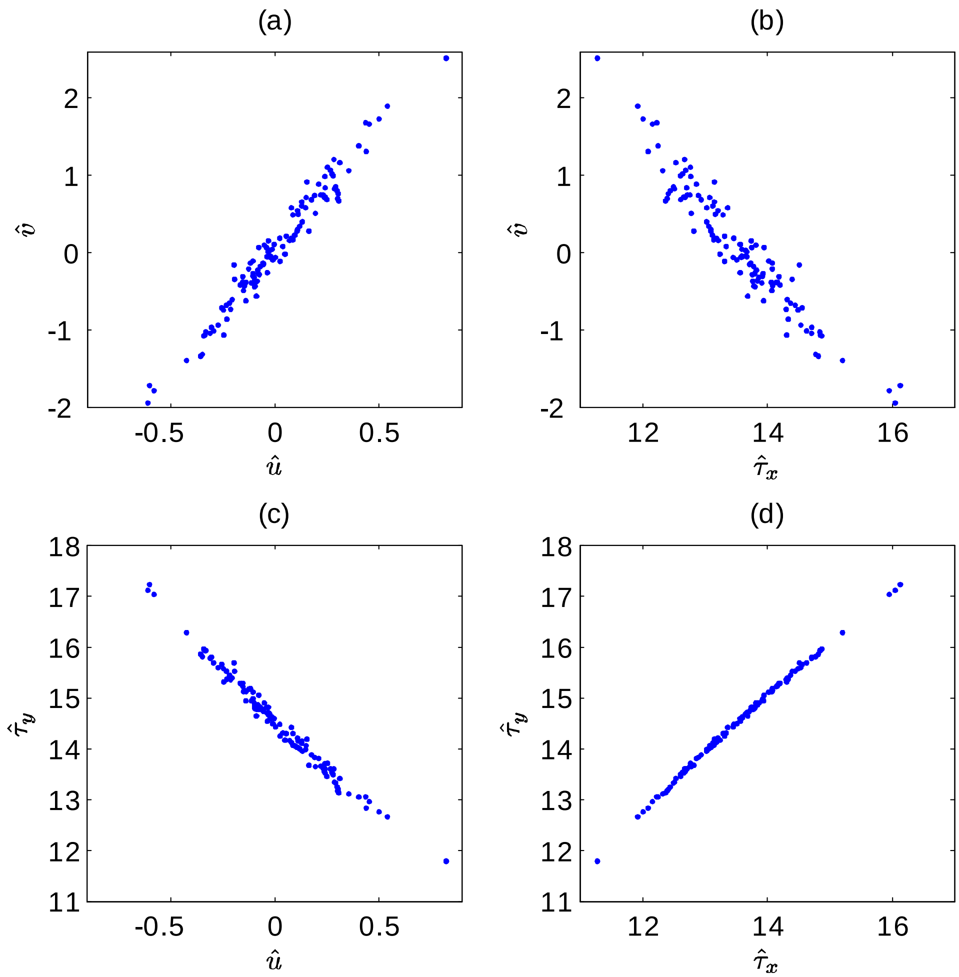

Figure 10: Zurich Growth Data Example. Individual random-effects estimators obtained by warped functional regression.

but since the first pc of the ys confounds amplitude and phase variability, it is not clear what proportion of the correlation is due to amplitude and what proportion is due to phase variability. In contrast, warped regression, while also showing a correlation between timing and amplitude of the peaks (Figures 10(b) and 10(c)), with late peaks being generally lower, does not confound the sources of variability.

Part VI

Cubic Hermite Splines

11 Interpolation on a single interval

11.1 Interpolation on the interval $[0, 1]$

Given values of the function to be interpolated, f_0 at $t = 0$ and f_1 at $t = 1$, and values of the derivatives at those points, d_0 at $t = 0$ and d_1 at $t = 1$, then

$$f(t) = h_{00}(t)f_0 + h_{10}(t)d_0 + h_{01}(t)f_1 + h_{11}(t)d_1$$

with

$$\begin{aligned}
h_{00}(t) &= 2t^3 - 3t^2 + 1 = (1 + 2t)(1 - t)^2 \\
h_{10}(t) &= t^3 - 2t^2 + t = t(1 - t)^2 \\
h_{01}(t) &= -2t^3 + 3t^2 = t^2(3 - 2t) = h_{00}(1 - t) \\
h_{11}(t) &= t^3 - t^2 = t^2(t - 1) = -h_{10}(1 - t)
\end{aligned}$$

satisfies

$$\begin{aligned}
f(0) &= f_0, \quad f(1) = f_1, \\
f'_+(0) &= d_0, \quad f'_-(1) = d_1.
\end{aligned}$$

Since $f(t)$ is a polynomial of degree 3 (i.e. has 4 free coefficients) and satisfies these four conditions, it's the *only* cubic polynomial that satisfies these conditions.

11.2 Interpolation on a general interval $[x_k, x_{k+1}]$

If we are now given values of the function f_k at $t = x_k$ and f_{k+1} at $t = x_{k+1}$, and values of the derivatives at those points, d_k at $t = x_k$ and d_{k+1} at $t = x_{k+1}$, then

$$f(t) = h_{00}\left(\frac{t - x_k}{s_k}\right) f_k + h_{10}\left(\frac{t - x_k}{s_k}\right) s_k d_k + h_{01}\left(\frac{t - x_k}{s_k}\right) f_{k+1} + h_{11}\left(\frac{t - x_k}{s_k}\right) s_k d_{k+1}$$

with $s_k = x_{k+1} - x_k$ and $h_{ij}(t)$ as before, satisfies

$$\begin{aligned} f(x_k) &= f_k, \;\; f(x_{k+1}) = f_{k+1}, \\ f'_+(x_k) &= d_k, \;\; f'_-(x_{k+1}) = d_{k+1}. \end{aligned}$$

Again, this cubic polynomial is unique, subject to these four conditions.

12 Interpolating a data set

Suppose we have p points in some interval $[a, b]$, $a < x_1 < x_2 < \cdots < x_p < b$, and corresponding values of the function, f_k, and the derivative, d_k, at each point x_k. Let's define $x_0 = a$, $x_{p+1} = b$ and f_0, f_{p+1}, d_0 and d_{p+1} accordingly. Then the piecewise cubic interpolant is going to be

$$\begin{aligned} f(t) &= \sum_{k=0}^{p} \Bigg\{ h_{00}\left(\frac{t-x_k}{s_k}\right) f_k + h_{10}\left(\frac{t-x_k}{s_k}\right) s_k d_k \\ &\quad + h_{01}\left(\frac{t-x_k}{s_k}\right) f_{k+1} + h_{11}\left(\frac{t-x_k}{s_k}\right) s_k d_{k+1} \Bigg\} \\ &= \sum_{k=0}^{p} h_{00}\left(\frac{t-x_k}{s_k}\right) f_k + \sum_{k=0}^{p} h_{10}\left(\frac{t-x_k}{s_k}\right) s_k d_k \\ &\quad + \sum_{k=1}^{p+1} h_{01}\left(\frac{t-x_{k-1}}{s_{k-1}}\right) f_k + \sum_{k=1}^{p+1} h_{11}\left(\frac{t-x_{k-1}}{s_{k-1}}\right) s_{k-1} d_k \end{aligned}$$

$$= h_{00}\left(\frac{t-x_0}{s_0}\right) f_0 + h_{10}\left(\frac{t-x_0}{s_0}\right) s_0 d_0 + \sum_{k=1}^{p} \left\{ h_{00}\left(\frac{t-x_k}{s_k}\right) + h_{01}\left(\frac{t-x_{k-1}}{s_{k-1}}\right) \right\} f_k$$

$$+ \sum_{k=1}^{p} \left\{ h_{10}\left(\frac{t-x_k}{s_k}\right) s_k + h_{11}\left(\frac{t-x_{k-1}}{s_{k-1}}\right) s_{k-1} \right\} d_k + h_{01}\left(\frac{t-x_p}{s_p}\right) f_{p+1} + h_{11}\left(\frac{t-x_p}{s_p}\right) s_p d_{p+1}$$

with $s_k = x_{k+1} - x_k$ as before, and we are defining $h_{ij}(t) = 0$ for $t \notin [0, 1]$.

Then we can introduce the basis functions

$$\phi_k(t) = \begin{cases} h_{00}\left(\frac{t-x_0}{s_0}\right) & \text{if } k = 0 \\ h_{00}\left(\frac{t-x_k}{s_k}\right) + h_{01}\left(\frac{t-x_{k-1}}{s_{k-1}}\right) & \text{if } k = 1, \ldots, p \\ h_{01}\left(\frac{t-x_p}{s_p}\right) & \text{if } k = p+1 \end{cases}$$

and

$$\psi_k(t) = \begin{cases} h_{10}\left(\frac{t-x_0}{s_0}\right)s_0 & \text{if } k=0 \\ h_{10}\left(\frac{t-x_k}{s_k}\right)s_k + h_{11}\left(\frac{t-x_{k-1}}{s_{k-1}}\right)s_{k-1} & \text{if } k=1,\ldots,p \\ h_{11}\left(\frac{t-x_p}{s_p}\right)s_p & \text{if } k=p+1, \end{cases}$$

and we have

$$\boxed{f(t) = \sum_{k=0}^{p+1}\phi_k(t)f_k + \sum_{k=0}^{p+1}\psi_k(t)d_k}$$

Since $h_{01}(t) = h_{00}(1-t)$ and $h_{11}(t) = -h_{10}(1-t)$,

$$\phi_k(t) = \begin{cases} h_{00}\left(\frac{t-x_0}{s_0}\right) & \text{if } k=0 \\ h_{00}\left(\frac{t-x_k}{s_k}\right) + h_{00}\left(\frac{x_k-t}{s_{k-1}}\right) & \text{if } k=1,\ldots,p \\ h_{00}\left(\frac{x_{p+1}-t}{s_p}\right) & \text{if } k=p+1 \end{cases}$$

and

$$\psi_k(t) = \begin{cases} h_{10}\left(\frac{t-x_0}{s_0}\right)s_0 & \text{if } k=0 \\ h_{10}\left(\frac{t-x_k}{s_k}\right)s_k - h_{10}\left(\frac{x_k-t}{s_{k-1}}\right)s_{k-1} & \text{if } k=1,\ldots,p \\ -h_{10}\left(\frac{x_{p+1}-t}{s_p}\right)s_p & \text{if } k=p+1. \end{cases}$$

For $k = 1,\ldots,p$ we have

$$\phi_k(t) = \begin{cases} 0 & \text{if } t < x_{k-1} \text{ or } t > x_{k+1} \\ h_{00}\left(\frac{x_k-t}{s_{k-1}}\right) & \text{if } x_{k-1} \le t \le x_k \\ h_{00}\left(\frac{t-x_k}{s_k}\right) & \text{if } x_k \le t \le x_{k+1} \end{cases}$$

and

$$\psi_k(t) = \begin{cases} 0 & \text{if } t < x_{k-1} \text{ or } t > x_{k+1} \\ -h_{10}\left(\frac{x_k-t}{s_{k-1}}\right)s_{k-1} & \text{if } x_{k-1} \le t \le x_k \\ h_{10}\left(\frac{t-x_k}{s_k}\right)s_k & \text{if } x_k \le t \le x_{k+1}. \end{cases}$$

Since $h_{00}(0) = 1$, $h_{00}(1) = 0$ and $h'_{00}(0) = h'_{00}(1) = 0$, the ϕ_ks are continuous and differentiable everywhere as functions of t, with $\phi_k(x_{k-1}) = \phi_k(x_{k+1}) = 0$, $\phi_k(x_k) = 1$ and $\phi'_k(x_{k-1}) = \phi'_k(x_k) = \phi'_k(x_{k+1}) = 0$. Similarly, since $h_{10}(0) = h_{10}(1) = 0$, $h'_{10}(0) = 1$ and $h'_{10}(1) = 0$, the ψ_ks are also continuous and differentiable everywhere as functions of t, with $\psi_k(x_{k-1}) = \psi_k(x_k) = \psi_k(x_{k+1}) = 0$, $\psi'_k(x_{k-1}) = \psi'_k(x_{k+1}) = 0$ and $\psi'_k(x_k) = 1$.

For the “border” basis functions we have

$$\phi_0(t) = \begin{cases} 0 & \text{if } t < a \text{ or } t > x_1 \\ h_{00}\left(\frac{t-a}{x_1-a}\right) & \text{if } a \le t \le x_1 \end{cases}$$

which is discontinuous only at $t = a$, with $\phi_0(a) = 1$, $\phi_0(x_1) = 0$, and $(\phi_0)'_+(a) = \phi_0'(x_1) = 0$;

$$\phi_{p+1}(t) = \begin{cases} 0 & \text{if } t < x_p \text{ or } t > b \\ h_{00}\left(\frac{b-t}{b-x_p}\right) & \text{if } x_p \le t \le b \end{cases}$$

which is discontinuous only at $t = b$, with $\phi_{p+1}(x_p) = 0$, $\phi_{p+1}(b) = 1$, and $\phi_{p+1}'(x_p) = \left(\phi_{p+1}\right)'_-(b) = 0$;

$$\psi_0(t) = \begin{cases} 0 & \text{if } t < a \text{ or } t > x_1 \\ h_{10}\left(\frac{t-a}{x_1-a}\right) s_0 & \text{if } a \le t \le x_1 \end{cases}$$

which is continuous everywhere, with $\psi_0(a) = \psi_0(x_1) = 0$, $(\psi_0)'_+(a) = 1$, and $\psi_0'(x_1) = 0$; and

$$\psi_{p+1}(t) = \begin{cases} 0 & \text{if } t < x_p \text{ or } t > b \\ -h_{10}\left(\frac{b-t}{b-x_p}\right) s_p & \text{if } x_p \le t \le b \end{cases}$$

which is continuous everywhere, with $\psi_{p+1}(x_p) = \psi_{p+1}(b) = 0$, $\psi_{p+1}'(x_p) = 0$, and $\left(\psi_{p+1}\right)'_-(b) = 1$.

13 Monotone interpolation

Given $a = x_0 < x_1 < \cdots < x_p < x_{p+1} = b$ and $f_0 < \cdots < f_{p+1}$, with d_ks unspecified, it's always possible to find d_ks such that the resulting $f(t)$ is strictly increasing (Fritsch and Carlson, 1980). Let

$$\Delta_k = \frac{f_{k+1} - f_k}{x_{k+1} - x_k}, \quad \alpha_k = \frac{d_k}{\Delta_k}, \quad \beta_k = \frac{d_{k+1}}{\Delta_k}.$$

Then $f(t)$ is monotone in $[x_k, x_{k+1}]$ if and only if:

1. $\alpha_k + \beta_k - 2 \le 0$ and $\operatorname{sign}(d_k) = \operatorname{sign}(d_{k+1}) = \operatorname{sign}(\Delta_k)$; or

2. $\alpha_k + \beta_k - 2 > 0$, $\operatorname{sign}(d_k) = \operatorname{sign}(d_{k+1}) = \operatorname{sign}(\Delta_k)$, and:

   (a) $2\alpha_k + \beta_k - 3 \le 0$, or

(b) $\alpha_k + 2\beta_k - 3 \leq 0$, or

(c) $\alpha_k - (2\alpha_k + \beta_k - 3)^2/\{3(\alpha_k + \beta_k - 2)\} \geq 0$.

The condition

$$\sqrt{\alpha_k^2 + \beta_k^2} \leq 3$$

implies either 1 or 2(a)–2(c) above, so it's sufficient to guarantee monotonicity. This motivates the following algorithm for constructing the d_ks:

1. Initialize the derivatives $\{d_k\}$ so that $\text{sign}(d_k) = \text{sign}(d_{k+1}) = \text{sign}(\Delta_k)$. For instance,

$$d_0 = \Delta_0, \ \ d_k = \frac{\Delta_{k-1} + \Delta_k}{2} \text{ for } k = 1, \ldots, p, \ \ d_{p+1} = \Delta_p.$$

2. For $k = 0, \ldots, p$:

   (a) If $\sqrt{\alpha_k^2 + \beta_k^2} \leq 3$ the interpolant will be monotone in $[x_k, x_{k+1}]$; go to next k.

   (b) If $\sqrt{\alpha_k^2 + \beta_k^2} > 3$, let $\tau_k = 3/\sqrt{\alpha_k^2 + \beta_k^2}$, $\alpha_k^* = \tau_k \alpha_k$, and $\beta_k^* = \tau_k \beta_k$; set

$$d_k = \alpha_k^* \Delta_k, \ \ d_{k+1} = \beta_k^* \Delta_k.$$

   The interpolant will be monotone in $[x_k, x_{k+1}]$; go to next k.

The algorithm may change the value of each d_k at most twice from its initial value: first when the interval $[x_{k-1}, x_k]$ is considered and again when the interval $[x_k, x_{k+1}]$ is considered. But since $0 \leq \alpha_k^* \leq \alpha_k$ and $0 \leq \beta_k^* \leq \beta_k$, the modification of d_k for $[x_k, x_{k+1}]$ will maintain the monotonicity condition on $[x_{k-1}, x_k]$; see comments on p. 241 of Fritsch and Carlson (1980).

14 References

Fritsch, F.N. and Carlson, R.E. (1980). Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* **17** 238–246.

Geyer, C. J. (1994). On the asymptotics of constrained M-estimators. *Ann. Statist.* **22** 1993–2010.

Kneip, A., and Ramsay, J. O. (2008). Combining registration and fitting for functional models. *Journal of the American Statistical Association* **103** 1155–1165.

Magnus, J.R. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics (Second Edition)*. Wiley, New York.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.

Prader, A., Largo, R. H., Molinari, L., & Issler, C. (1989). Physical growth of Swiss children from birth to 20 years of age. First Zurich longitudinal study of growth and development. *Helvetica paediatrica acta. Supplementum* **52** 1–125.

Rockafellar, R. & Wets, R. (1998). *Variational Analysis*. New York: Springer.

Sangalli, L.M., Secchi, P., Vantini, S., and Vitelli, V. (2010). K-mean alignment for curve clustering. *Computational Statistics and Data Analysis* **54** 1219–1233.