Curve Registration

Anthony Ebert

14 February 2019

Motivation

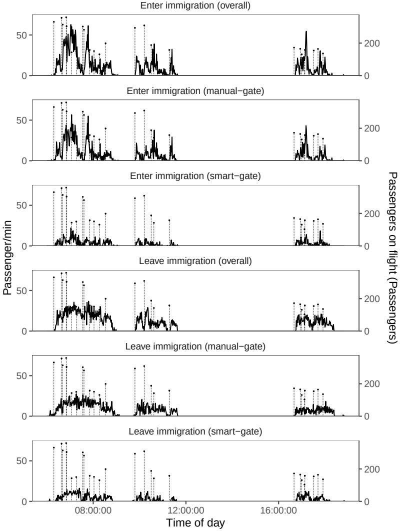

Figure 1: Passenger flow counts.

Introduction

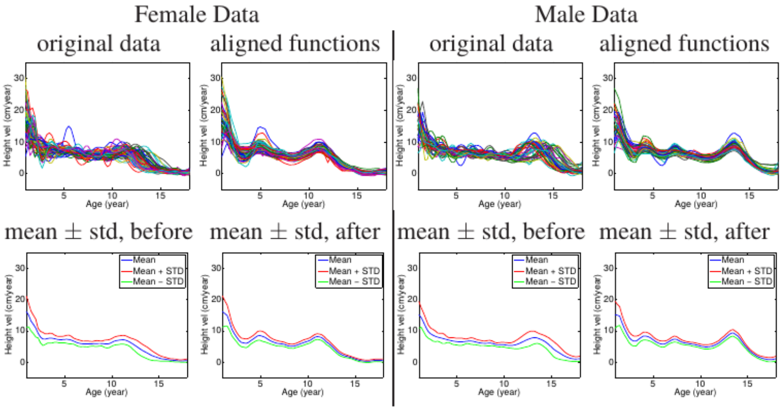

Figure 2: Analysis of growth data, before and after curve registration (Srivastava et al. 2011).

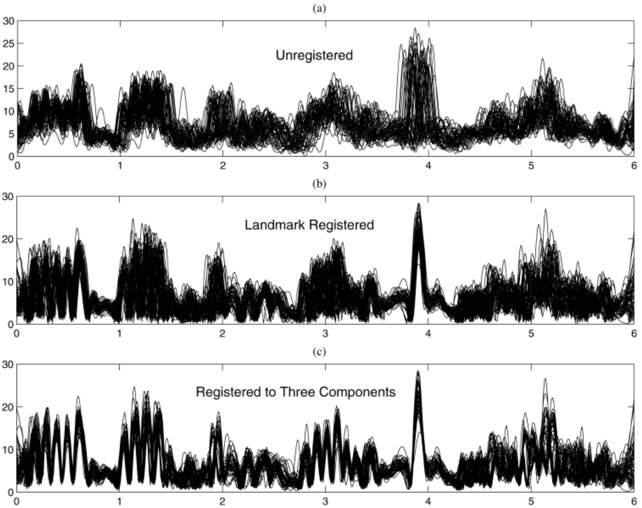

Figure 3: 50 curves before and after registration (Kneip and Ramsay 2008).

Functional data analysis

Functional data analysis (FDA) is the analysis of data generated by curves or functions (Hsing and Eubank 2015). Special-cases of functional datasets include longitudinal data and time series. However, the field of FDA is much broader than this (Wang, Chiou, and Müller 2016). Functional random variables take values in a function space $\mathrm{f} \in \mathcal{F}$ with associated probability measure p (see Delaigle and Hall (2010)). In practice, the functional random variables are observed with error on a countable subset of the domain of the function. This is what we refer to as an empirical functional random variable (EFRV). An EFRV f is itself a set of pairs $f_1, f_2, \cdots$ of the form $f_i = (t_i^f, y_i^f)$ representing sampling location t_i^f and corresponding function output y_i^f. Therefore f is of the form $\{(t_1^f, y_1^f), \cdots, (t_j^f, y_j^f), \cdots, (t_n^f, y_n^f)\}$. We append the superscript f to keep track of which sampling points and functional outputs refer to which EFRV.

We standardise the domain of $\mathsf{f}^* \in \mathbb{R}^{[0,T]}$ from $[0, T]$ to $[0, 1]$ by setting $\mathsf{f}(t) = \mathsf{f}^*(t \times T)$. The idea of curve registration is to align elements of $\mathcal{F}$ with warping functions $\gamma^{\mathsf{f}} : [0,1] \mapsto [0,1]$, such that elements of the set $\mathcal{G} := \{\mathsf{f} \circ \gamma^{\mathsf{f}} | \mathsf{f} \in \mathcal{F}\}$ have aligned features according to some criteria. We adopt the elastic functions approach of Srivastava et al. (2011), who align f to another functional random variable g using the Fisher-Rao metric:

$$d_{\mathsf{FR}}(\mathsf{f}, \mathsf{g}) = \int [q^{\mathsf{f}}(t) - q^{\mathsf{g}}(t)]^2 \mathrm{d}t, \tag{1}$$

where $q^{\mathsf{f}}(t) = \mathrm{sign}(\mathsf{f}'(t)) \times \sqrt{|\mathsf{f}'(t)|}$. Curve f is aligned to g by defining a warping function $\gamma \in \Gamma$ where $\Gamma = \{\gamma \in [0,1]^{[0,1]} | \gamma \text{ is invertible and } \gamma(0) = 0\}$ to minimise $d_{\mathsf{FR}}(\mathsf{f} \circ \gamma, \mathsf{g})$.

Curve f is aligned to g by defining a warping function $\gamma \in \Gamma$ where $\Gamma = \{\gamma \in [0,1]^{[0,1]} | \gamma \text{ is invertible and } \gamma(0) = 0\}$ to minimise $d_{\mathsf{FR}}(\mathsf{f} \circ \gamma, \mathrm{g})$. One advantage of using d_{FR} for curve registration is that:

$$d_{\mathsf{FR}}(\mathsf{f}, \mathrm{g}) = d_{\mathsf{FR}}(\mathsf{f} \circ \gamma, \mathrm{g} \circ \gamma),$$

in other words, the Fisher-Rao metric is invariant to shared warpings.

This implies that the discrepancy of amplitudes (amplitude distance) between f and g, defined as

$$d_{\mathsf{amp}}(\mathsf{f}, \mathsf{g}) := \inf_{\gamma \in \Gamma} d_{\mathsf{FR}}(\mathsf{f} \circ \gamma, \mathsf{g}),$$

is symmetric (Srivastava et al. 2011).

Approximate Bayesian computation

The FR metric is a dissimilarity, another dissimilarity which we use is the estimator for maximum mean discrepancy (d_{MMD}), developed by Gretton et al. (2007)

The definition of $d_{\mathrm{MMD}}(f, g)$ is as follows:

$$\begin{aligned} d_{\mathrm{MMD}}(f, g) = & \frac{1}{m^2} \sum_{j=1}^{m} \sum_{j'=1}^{m} k(f_j, f_{j'}) + \\ & \frac{1}{n^2} \sum_{j=1}^{n} \sum_{j'=1}^{n} k(g_j, g_{j'}) \\ & - \frac{2}{mn} \sum_{j=1}^{m} \sum_{j'=1}^{n} k(f_j, g_{j'}), \end{aligned} \tag{2}$$

where m is the cardinality of f, n is the cardinality of g and k is a kernel function.

A common choice of kernel function is the Gaussian kernel, $k(f_j, g_{j'}) = \exp\left[-0.5\sqrt{(t_j^f - t_{j'}^g)^T S^{-1}(y_j^f - y_{j'}^g)}\right]$, where S is a fixed tuning covariance matrix. Gretton et al. (2012) showed that d_{MMD} is equivalent to a kernel-smoothed L2 norm between EFRVs. We can, therefore, use $\hat{\rho}_{\mathrm{MMD}}$ as a dissimilarity on EFRVs rather than probability measures.

Gaussian peak shift

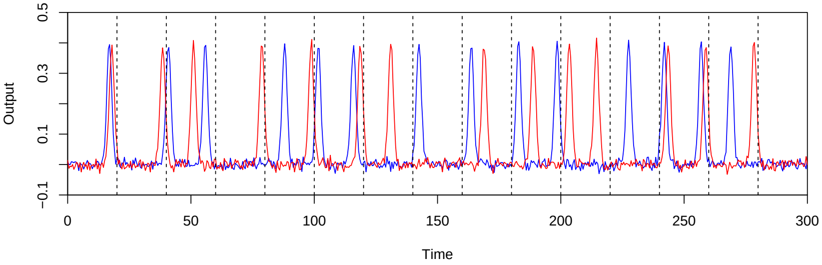

Figure 4: Example of model realisations from the Gaussian peak shift model with $\sigma_b = 5, \sigma_\phi = 1$, and $\sigma_\epsilon = 0.01$. The dotted lines are the μ_u values. The distribution of peak shifts is controlled by σ_b, the peak widths are controlled by σ_ϕ, and the noise is controlled by σ_ϵ.

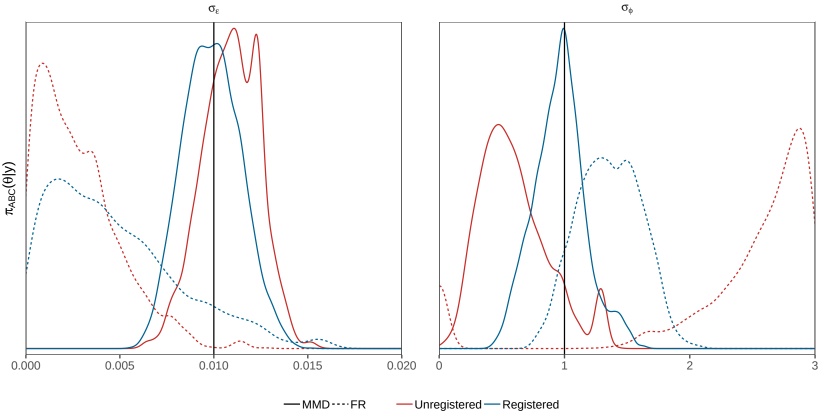

Figure 5: Density plots of posterior samples arising from the replenishment ABC sampler (Drovandi and Pettitt 2011) for the Gaussian peak shift example. Distances shown include MMD (Maximum mean discrepancy) and FR (Fisher-Rao) in on both registered and unregistered data. Vertical black solid lines represent true values.

Passenger processing

Algorithm 3 Airport loss function

function $(\theta, f^a, f^z, \text{type})$
  $\mathbf{a}_\theta^{\text{imm}} \sim p_a(\cdot|\mu, \nu)$
  $f_\theta^a = \text{hist}(\mathbf{a}_\theta^{\text{imm}})$
  $\breve{\mathbf{a}}_\theta^{\text{imm}} = \mathbf{a}_\theta^{\text{imm}}$
  if type = registered or corrected **then**
    $\gamma = \arg\inf_{\gamma\in\Gamma} d_{\text{FR}}(f^a, f_\theta^a \circ \gamma)$
    $d_\theta^a = d(f^a, f_\theta^a \circ \gamma)$
    if type = corrected **then**
      $\breve{\mathbf{a}}_\theta^{\text{imm}} = \text{warp}(\mathbf{a}_\theta^{\text{imm}}, \gamma)$
    end if
  else
    $d_\theta^a = d(f^a, f_\theta^a)$
  end if
  $\mathbf{z}_\theta^{\text{imm}} \sim p_a(\cdot|\breve{\mathbf{a}}_\theta^{\text{imm}}, \lambda, K)$
  $f_\theta^z = \text{hist}(\mathbf{z}_\theta^{\text{imm}})$
  $d_\theta = d_\theta^a + d(f^z, f_\theta^z)$
  return d_θ
end function

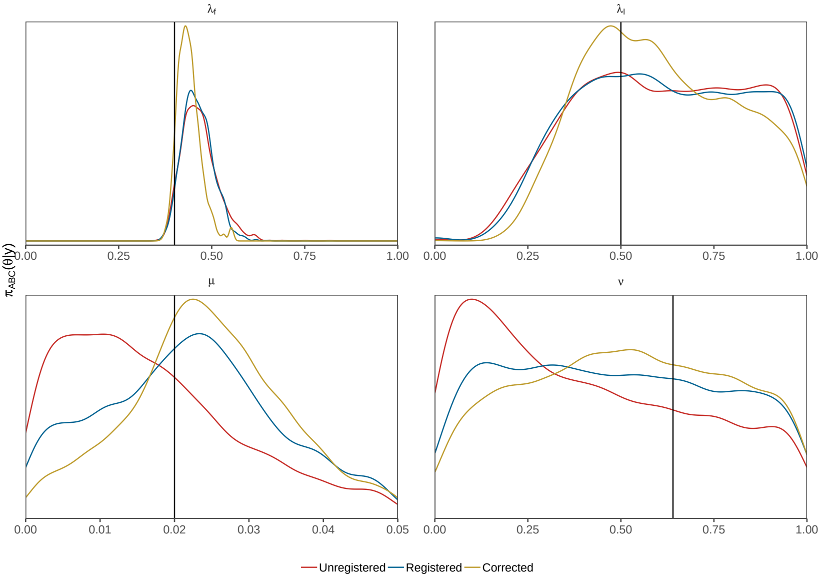

Figure 6: Density plots of posterior samples using the airport loss function (Algorithm 3) for unregistered, registered and corrected types. The solid vertical black line represents the true value.