Research Article

Statistics in Medicine

Received XXXX

(www.interscience.wiley.com) DOI: 10.1002/sim.0000

Supplementary materials for: "Discrimination surfaces with application to region-specific brain asymmetry analysis"

Gabriel Martos[a] and Miguel de Carvalho[b]*

1. Supplementary technical details and proof of Lemma 2

We start with some comments on monotonicity of the $F_{\boldsymbol{t}}^{-1}(p) = \inf\{y : F_{\boldsymbol{t}}(y) \geqslant p\}$. Let $0 \leqslant a_0 \leqslant a_1 < b \leqslant 1$. Note that since $F_{(a,\cdot)}(\cdot)$ is nondecreasing it follows that $F_{(a_0,b)}(y) \leqslant F_{(a_1,b)}(y)$, for all $y \in S$, and so

$$F_{(a_0,b)}^{-1}(y) = \inf\{y : F_{(a_0,b)}(y) \geqslant p\} \leqslant \inf\{y : F_{(a_1,b)}(y) \geqslant p\} = F_{(a_1,b)}^{-1}(y). \tag{1}$$

Thus, Equation (1) implies that $F_{(a,\cdot)}^{-1}(\cdot)$ is nondecreasing; similar arguments can be used to show that $F_{(\cdot,b)}^{-1}(\cdot)$ is nonincreasing.

We now prove Lemma 2. Below, $\widehat{F}_{\boldsymbol{t}}^{-1}(p) = \inf\{y : \widehat{F}_{\boldsymbol{t}}(y) \geqslant p\}$ are the marginal empirical quantiles, with $\widehat{F}_{\boldsymbol{t}}(y) = n^{-1}\sum_{i=1}^{n} I\{Y_i(\boldsymbol{t}) \leqslant y\}$ denoting the marginal empirical distribution function. In addition, $\mathscr{Y}$, is the space of all non-negative and differentiable random functions on T, such that $\partial Y(\boldsymbol{t})/\partial a \leqslant 0$ and $\partial Y(\boldsymbol{t})/\partial b \geqslant 0$, and which are supported over $S = [0, M]$, for some $M > 0$, for every $\boldsymbol{t} \in T = \{(a, b) \in [0, 1]^2 : 0 \leqslant a < b \leqslant 1\}$.

Proof of Lemma 2. Since we are assuming that $F_{\boldsymbol{t}}(y)$ is strictly increasing on $[0, M]$ it follows that $F_{\boldsymbol{t}}^{-1}(p)$ is continuous for all $p \in [F_{\boldsymbol{t}}^{-1}(0), F_{\boldsymbol{t}}^{-1}(1)] = [0, 1]$, by keeping in mind [1, Problem 45] and the assumption that random curves in $\mathscr{Y}$ are supported over $S = [0, M]$, for some $M > 0$.

Our line of attack is now similar to that of [2, p. 62]. Let $C_{i,j,k} = [p_i, p_{i+1}] \times [a_j, a_{j+1}] \times [b_k, b_{k+1}]$, with

$$0 = p_0 < p_1 < \cdots < p_{I-1} < p_I = 1, \quad 0 = a_0 < a_1 < \cdots < a_{J-1} < a_J = 1, \quad 0 = b_0 < b_1 < \cdots < b_{K-1} < b_K = 1,$$

be such

$$|F(p_{i+1}, a_{j+1}, b_k) - F(p_i, a_j, b_{k+1})| < \varepsilon,$$

for a given $\varepsilon > 0$, for $i = 0, \ldots, I$, $j = 0, \ldots, J$, and $k = 0, \ldots, K$. By the monotonicity properties of $F_{\boldsymbol{t}}^{-1}(p)$ and $\widehat{F}_{\boldsymbol{t}}^{-1}(p)$ on $[0, 1]^3$ (namely: $F_{(\cdot,\cdot)}^{-1}(p)$ and $F_{(a,\cdot)}^{-1}(\cdot)$ are nondecreasing, $F_{(\cdot,b)}^{-1}(\cdot)$ is nonincreasing, and analogous properties hold for

[a] *Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina.* [b] *School of Mathematics, The University of Edinburgh, UK.*

* *Correspondence to: M. de Carvalho. School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD. Email: miguel.decarvalho@ed.ac.uk.*

Prepared using simauth.cls *[Version: 2010/03/10 v3.00]*

$\widehat{F}_{\boldsymbol{t}}^{-1}(p))$, it follows that

$$\begin{aligned}\nabla_n \equiv \sup_{(p,\boldsymbol{t})\in[0,1]\times T} |\widehat{F}_{\boldsymbol{t}}^{-1}(p) - F_{\boldsymbol{t}}^{-1}(p)| &\leqslant \sup_{(p,\boldsymbol{t})\in[0,1]\times[0,1]^2} |\widehat{F}_{\boldsymbol{t}}^{-1}(p) - F_{\boldsymbol{t}}^{-1}(p)| \\ &= \max_{i,j,k} \sup_{(p,\boldsymbol{t})\in C_{i,j,k}} |\widehat{F}_{\boldsymbol{t}}^{-1}(p) - F_{\boldsymbol{t}}^{-1}(p)| \\ &\leqslant \max_{i,j,k} \max\{|\widehat{F}^{-1}(p_{i+1}, a_{j+1}, b_k) - F^{-1}(p_i, a_j, b_{k+1})|, \\ &\qquad |F^{-1}(p_{i+1}, a_{j+1}, b_k) - \widehat{F}^{-1}(p_i, a_j, b_{k+1})|\},\end{aligned}$$

and thus by taking the limit and using standard arguments on pointwise convergence of $F_{\boldsymbol{t}}^{-1}(p)$ [3, p. 305], it follows that

$$\limsup_{n\to\infty} \nabla_n = \max_{i,j,k}\{|F^{-1}(p_{i+1}, a_{j+1}, b_k) - F^{-1}(p_i, a_j, b_{k+1})|\} < \varepsilon, \quad \text{a.s.}$$

$\square$

2. Transforming brain raw data into functional data

Most functional data analysis approaches for preprocessing raw data suggest proceeding as follows: Choose an orthogonal basis of functions $\Phi = \{\phi_1, \ldots, \phi_N\}$, where each ϕ_i belongs to a general function space $\mathcal{H}$, and then represent each functional datum by means of a linear combination in the $\mathrm{Span}(\Phi)$ [4, 5]. A usual choice is to consider $\mathcal{H}$ as a Reproducing Kernel Hilbert Space (RKHS) of functions [6]. In this case, the elements in the spanning set Φ are the eigenfunctions associated to the positive-definite and symmetric kernel function $K_h : A \times A \to \mathbb{R}$ that span $\mathcal{H}$, and with $h > 0$ being a bandwidth.

In our setting, $A = [0, 2\pi]$ and the functional representation problem can be framed as follows: We have $m = 500$ (x, z) coordinates for each of the $n = 68$ brain curves (cf Figure 2 (a) in the paper), and thus let $\{(\theta_j, r(\theta_j)) \in [0, 2\pi] \times \mathbb{R}\}_{j=1}^m$ be the discrete version of a brain curve $B \in \mathcal{H}$ represented in a polar coordinate system. The functional data estimator for each brain curve is obtained solving the following regularization problem:

$$\tilde{r}(\theta) := \arg\min_{g\in\mathcal{H}} \sum_{j=1}^m V(r(\theta_j), g(\theta_j))^2 + \gamma\Omega(g), \tag{2}$$

where V is a strictly convex functional with respect to the second argument, $\gamma > 0$ is a regularization parameter (chosen by cross–validation), and $\Omega(g)$ is a regularization term. By the Representer Theorem [7, Th. 5.2, p. 91] [8, Pr. 8, p. 51] the solution of the problem stated in Equation (2) exists, is unique, and admits a representation of the form

$$\tilde{r}(\theta) = \sum_{j=1}^m \alpha_j K_h(\theta, \theta_j). \tag{3}$$

where K_h is a kernel—say, $K_h(x, y) = \exp\{-(x - y)^2/2h\}$. In the particular case of a squared loss function $V(w, z) = (w - z)^2$ and considering $\Omega(g) = \int_0^{2\pi} g^2(\theta)\, \mathrm{d}\theta$, the coefficients of the linear combination in Equation (3) are obtained solving the following linear system

$$(\gamma m\mathbf{I} + \mathbf{K})\boldsymbol{\alpha} = \mathbf{y}, \tag{4}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)^{\mathrm{T}}$, $\mathbf{y} = (r(\theta_1), \ldots, r(\theta_m))^{\mathrm{T}}$, $\mathbf{I}$ is the identity matrix of order m, and $\mathbf{K}$ the is the Gram matrix with the kernel evaluations, $[\mathbf{K}]_{k,l} = K_h(\theta_k, \theta_l)$, for $k = 1, \ldots, m$ and $l = 1, \ldots, m$. In Figure 2, we show the raw data on the left and the estimated functional data obtained by solving the regularization problem in Equation (2). To obtain the brain curves in Figure 1 (b) in the paper, we use the Cartesian coordinate system, $\tilde{B}_i(\theta) = (\tilde{r}_i(\theta)\sin(\theta), \tilde{r}_i(\theta)\cos(\theta))$, where $\tilde{r}_i$ is the functional data estimator of the ith brain curve for $\theta \in [0, 2\pi]$ and $i = 1, \ldots, n$.

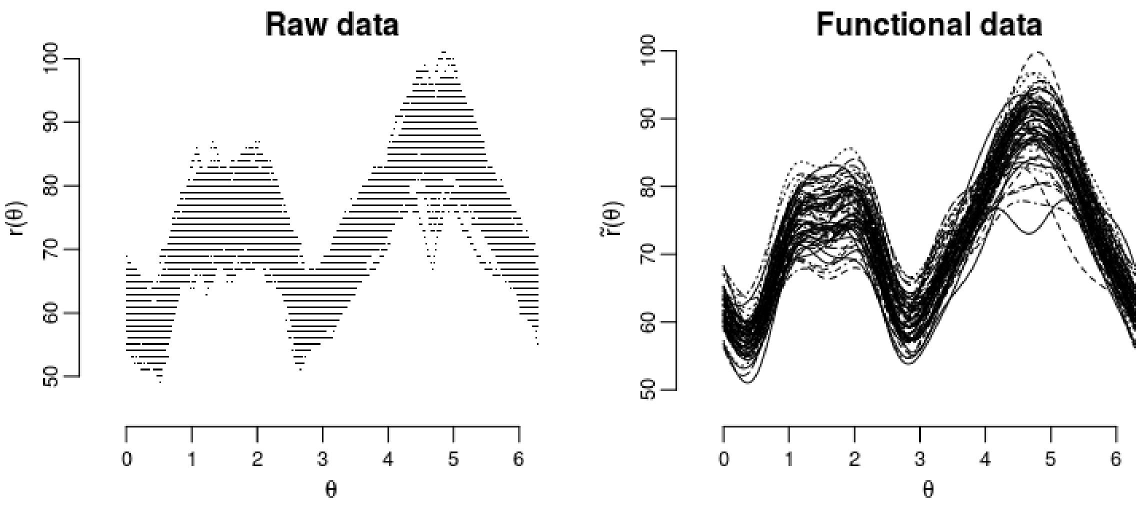

Figure 1. Raw data in polar coordinates (left) and brain curves obtained by solving the regularization problem in Equation (2) (right).

3. Supplementary empirical reports

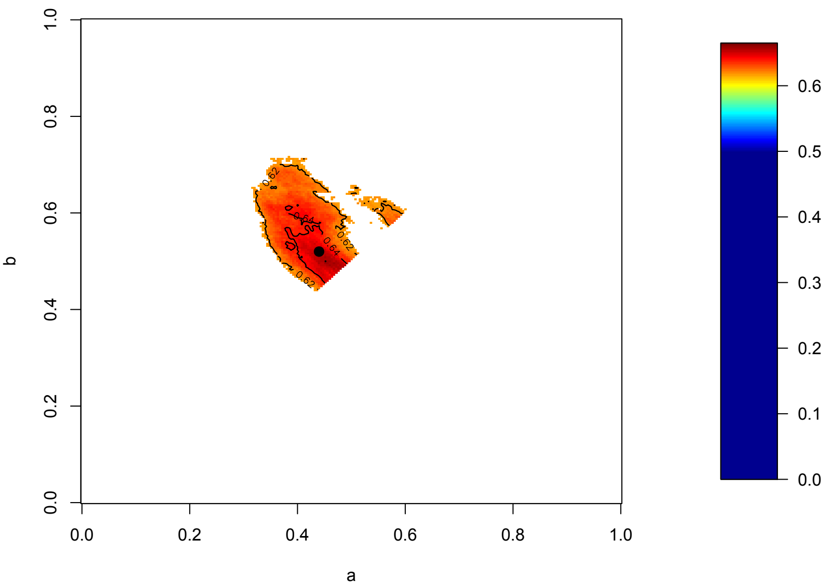

Figure 2. Region of T at which the null hypothesis $H_0 : \Lambda(\boldsymbol{t}) = 1/2$ is rejected in the data application in Section 5 in the paper.

4. `R` code

The following chunks of R code can used to replicate the numerical experiments introduced in Section 4 of the manuscript. Below we estimate the empirical discrimination surface and empirical discrimination level sets given a random sample of

(log)normal distributed asymmetry scores, as formally introduced in Example 1 of the manuscript.
We start by loading the relevant packages.

```
## Required packages:
list.of.packages <- c("akima", "fields", "MASS", "rgl", "Bolstad",
                      "hdrcde", "lattice")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[, "Package"])]
if(length(new.packages)) install.packages(new.packages) else {print('OK')}
 lapply(list.of.packages, require, character.only = TRUE)
```

We now setup the simulation settings and simulate data according to Scenario A.

```
## Localization parameters; T as formally defined in Equation (2).
T <- expand.grid(seq(0, 1, l = 100), seq(0, 1, l = 100))
T <- T[T[, 1] > T[, 2], ]
colnames(T) = c('b','a')

## Simulating random scores according to Scenario A
nD <- nnD <- 100
hat.Lambda <- c()
for(i in 1:dim(T)[1]) {
  b <- T[i, 1]
  a <- T[i, 2]
  alpha <- 1 - 10 * (b - 3 / 4)^2 - 10 * (a - 1 / 2)^2
  YD <-  rnorm(nD, alpha, 1)
  YnD <- rnorm(nnD, 0, 1)
  ans <- vapply(YD, function(YD) YD > YnD, logical(length(YnD)))
  hat.Lambda[i] <- sum(ans) / (nD * nnD)
}
```

We are now ready to fit the empirical discrimination surface underlying the simulated data. See Figure 3.

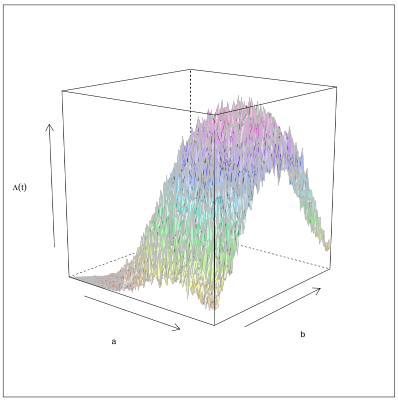

Figure 3. Empirical discrimination contour: Replicating Figure 3 1) in the manuscript.

The IMD estimates can be computed using the following code.

```
## IMD estimates
max(hat.Lambda)

## [1] 0.8312

T[which(max(hat.Lambda)==hat.Lambda, arr.ind = TRUE), ]

##              b         a
## 5277 0.7676768 0.5252525

## True: a* = 1/2 and b* = 3/4
```

The chunk of code below can be used to replicate the empirical discrimination contour from Figure 3 1) in the manuscript. See Figure 4.

```
## Empirical discrimination contour
resolution <- 0.01
interpol <- interp(x = hat.Lambda.t[, 2], y = hat.Lambda.t[, 1],
                   z = hat.Lambda.t[, 3], xo = seq(min(hat.Lambda.t[, 1]),
                   max(hat.Lambda.t[, 1]), by = resolution),
                   yo = seq(min(hat.Lambda.t[, 2]), max(hat.Lambda.t[, 2]),
                            by = resolution), duplicate = "mean")
filled.contour(interpol, color.palette = function(x)rev(heat.colors(x)),
               axes = TRUE, frame.plot = TRUE, xlab='a', ylab='b',
               plot.axes = {axis(1); axis(2);
               points(0.5, 0.75, pch = 19, lwd = 6)})
```

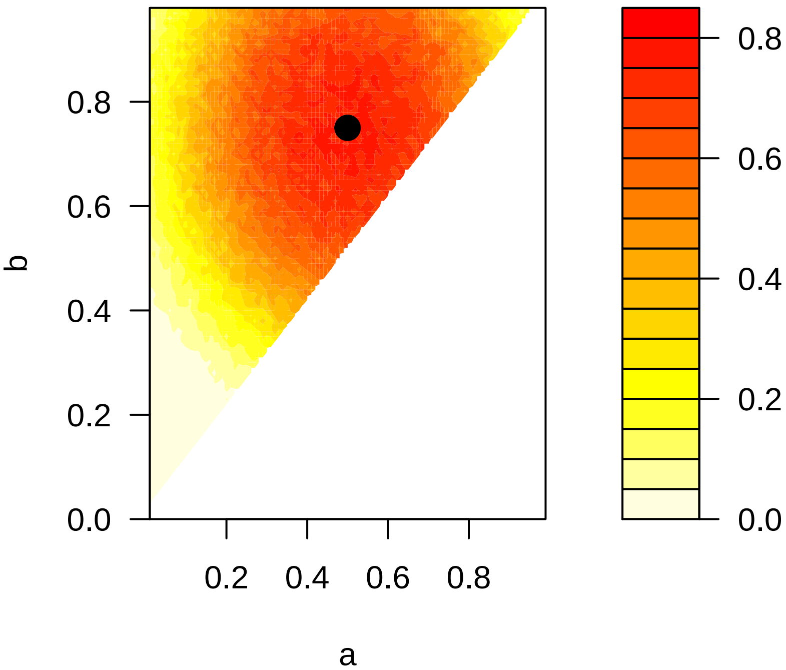

Figure 4. Empirical discrimination contour: Replicating Fig. 3 1) in the manuscript.

References

[1] Royden HL, Fitzpatrick PM. *Real Analysis*. 3rd ed. New York: Macmillan, 1988.

[2] Embrechts P, Klüppelberg C, Mikosch T. *Modelling Extremal Events for Insurance and Finance*. New York: Springer, 1997.

[3] van der Vaart AW. *Asymptotic Statistics*. Cambridge, UK: Cambridge University Press, 1998.

[4] Ramsay JO. *Functional Data Analysis*. Wiley: New York, 2006.

[5] Ferraty F, Vieu P. *Nonparametric Functional Data Analysis: Theory and Practice*. New York: Springer, 2006.

[6] Berlinet A, Thomas-Agnan C. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. New York: Springer, 2011.

[7] Kimeldorf G, Wahba G. Some results on Tchebycheffian spline functions. *Journal of Mathematical Analysis and Applications* 1971; **33**: 82–94.

[8] Cucker F, Smale S. On the mathematical foundations of learning. *Bulletin of the American Mathematical Society* 2002; **39**: 1–49.