Statistica Sinica: Supplement

AN OUTLYINGNESS MATRIX FOR MULTIVARIATE FUNCTIONAL DATA CLASSIFICATION

Wenlin Dai and Marc G. Genton[1]

CEMSE Division, King Abdullah University of Science and Technology,

Thuwal 23955-6900, Saudi Arabia.

This document serves as a supplement to the main manuscript. It mainly contains an illustrative example of functional directional outlyingness framework, two figures of real data, and technical proofs for the theoretical results.

Functional Directional Oultyingness

Figure S1 presents an example of functional directional outlyingness with a group of bivariate curves. In the graph on the left, the non-outlying curves are shown in black and outliers are shown in different colors. In the middle graph, the grey surface is a quadratic surface satisfying $\{(a, b, c) : c = a^2 + b^2; a,\ b,\ c > 0 \in \mathbb{R}\}$. Because the non-outlying curves and the two shifted outliers (blue and red curves) are mutually parallel, their mapped points, $(\mathbf{MO}^{\mathrm{T}}, \mathrm{FO})^{\mathrm{T}}$, fall exactly onto the grey quadratic surface. However, the two points corresponding with the shifted outliers are isolated from the cluster, making them easy to recognize. The right graph presents a scatter plot of $(\mathbf{MO}^{\mathrm{T}}, \mathrm{VO})^{\mathrm{T}}$, from which we can simply distinguish the cyan and purple points from the others by their VO values. The green point is not only isolated from the cluster, but it also has a larger VO, which coincides with its outlyingness for both scale and shape.

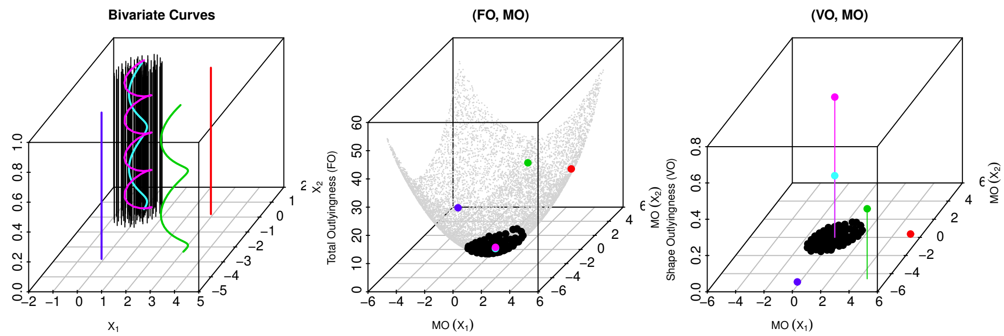

Figure S1: Left: 100 bivariate non-outlying curves with another five outliers; middle: functional directional outlyingness (FO) in relation to the mean directional outlyingness (**MO**); right: variation of directional outlyingness (VO) in relation to mean directional outlyingness (**MO**). The Mahalanobis depth is adopted for the calculation of directional outlyingness.

Phoneme Data

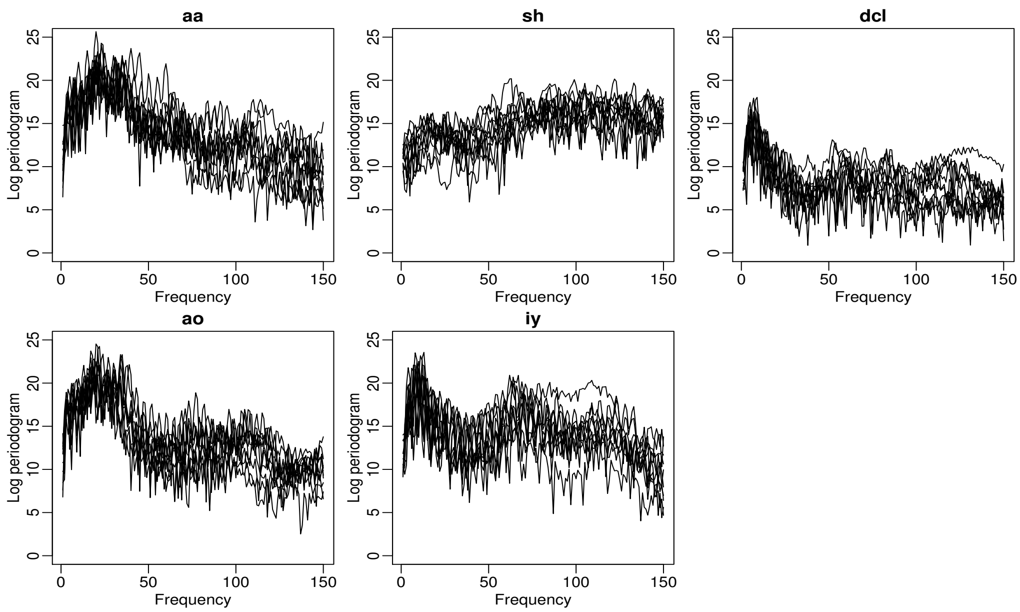

Figure S2: Ten samples for each of the five phonemes: “aa”, “ao”, “dcl”, “sh”, and “iy”.

Gesture Data

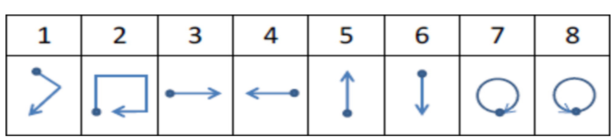

Figure S3: Gesture vocabulary. The dot denotes the start and the arrow the end of each of the eight gestures.

Proofs of Theoretical Results

Proofs of Theorem 1.

(i): Similarly, we decompose the total depth matrix into two parts: scale depth matrix and shape depth matrix:

$$\begin{aligned}
\mathbf{FOM}(\mathbf{X}, F_{\mathbf{X}}) &= \int_{\mathcal{I}} \mathbf{O}(\mathbf{X}(t), F_{\mathbf{X}(t)})\mathbf{O}^{\mathrm{T}}(\mathbf{X}(t), F_{\mathbf{X}(t)})w(t)\mathrm{d}t \\
&= \int_{\mathcal{I}} \left\{\mathbf{O} - \mathbf{MO} + \mathbf{MO}\right\}\left\{\mathbf{O} - \mathbf{MO} + \mathbf{MO}\right\}^{\mathrm{T}} w(t)\mathrm{d}t \\
&= \mathbf{VOM}(\mathbf{X}, F_{\mathbf{X}}) + \mathbf{MO}(\mathbf{X}, F_{\mathbf{X}})\mathbf{MO}^{\mathrm{T}}(\mathbf{X}, F_{\mathbf{X}}).
\end{aligned}$$

(ii) The result is straightforward using $\mathrm{tr}(\mathbf{a}\mathbf{a}^{\mathrm{T}}) = \mathbf{a}^{\mathrm{T}}\mathbf{a} = \|\mathbf{a}\|^2$ for any vector $\mathbf{a}$. □

Proof of Theorem 2.
To prove Theorem 2, we first prove the following results for directional outlyingness of point-wise data:

$$\mathbf{O}(\mathbf{A}_0\mathbf{X}(t) + \mathbf{b}, F_{\mathbf{A}_0\mathbf{X}(t)+\mathbf{b}}) = \mathbf{A}_0 \cdot \mathbf{O}(\mathbf{X}(t), F_{\mathbf{X}(t)}).$$

Proof: Since $d\left(\mathbf{X}(t), F_{\mathbf{X}(t)}\right)$ is a valid depth possessing the four popular properties stated in Definition 2.1 of Zuo and Serfling (2000), the affine invariance of $d\left(\mathbf{X}(t), F_{\mathbf{X}(t)}\right)$ indicates that $d\left(\mathbf{A}\mathbf{X}(t) + \mathbf{b}, F_{\mathbf{A}\mathbf{X}(t)+\mathbf{b}}\right) = d\left(\mathbf{X}(t), F_{\mathbf{X}(t)}\right)$. Consequently, we have

$$\|\mathbf{O}(\mathbf{A}_0\mathbf{X}(t) + \mathbf{b}, F_{\mathbf{A}_0\mathbf{X}(t)+\mathbf{b}})\| = \|\mathbf{O}(\mathbf{X}(t), F_{\mathbf{X}(t)})\|. \tag{S1}$$

For the directional part, $\mathbf{v}$, we have

$$\begin{aligned}
\mathbf{v}^*(t) &= \frac{\mathbf{A}_0\mathbf{X}(t) - \mathbf{A}_0\mathbf{Z}(t)}{\|\mathbf{A}_0\mathbf{X}(t) - \mathbf{A}_0\mathbf{Z}(t)\|} \\
&= \frac{\mathbf{A}_0\|\mathbf{X}(t) - \mathbf{Z}(t)\|}{\|\mathbf{A}_0\mathbf{X}(t) - \mathbf{A}_0\mathbf{Z}(t)\|} \cdot \frac{\mathbf{X}(t) - \mathbf{Z}(t)}{\|\mathbf{X}(t) - \mathbf{Z}(t)\|} \\
&= \mathbf{A}_0 \cdot \mathbf{v}(t).
\end{aligned} \tag{S2}$$

In the final step, we use $\|\mathbf{A}_0\mathbf{X}(t) - \mathbf{A}_0\mathbf{Z}(t)\| = \|\mathbf{X}(t) - \mathbf{Z}(t)\|$ since $\mathbf{A}_0$ is an orthogonal matrix. Then, based on (S1) and (S2), we get

$$\mathbf{O}(\mathbf{A}_0\mathbf{X}(t) + \mathbf{b}, F_{\mathbf{A}_0\mathbf{X}(t)+\mathbf{b}}) = \mathbf{A}_0 \cdot \mathbf{O}(\mathbf{X}(t), F_{\mathbf{X}(t)}). \tag{S3}$$

Since g is a one-to-one transformation on the interval $\mathcal{I}$, it is easy to show

$$\begin{aligned}
\mathbf{MO}\left(\mathbf{T}(\mathbf{X}), F_{\mathbf{T}(\mathbf{X})}\right) &= \mathbf{MO}\left(\mathbf{T}(\mathbf{X}_g), F_{\mathbf{T}(\mathbf{X}_g)}\right), \\
\mathbf{FOM}\left(\mathbf{T}(\mathbf{X}), F_{\mathbf{T}(\mathbf{X})}\right) &= \mathbf{FOM}\left(\mathbf{T}(\mathbf{X}_g), F_{\mathbf{T}(\mathbf{X}_g)}\right).
\end{aligned}$$

By (S3) and $f(t) > 0$, $\mathbf{O}(\mathbf{T}(\mathbf{X}(t)), F_{\mathbf{T}(\mathbf{X}(t))}) = \mathbf{A}_0\mathbf{O}(\mathbf{X}(t), F_{\mathbf{X}(t)})$ holds. Then, we have

$$\begin{aligned}
\mathbf{MO}(\mathbf{T}(\mathbf{X}), F_{\mathbf{T}(\mathbf{X})}) &= \{\lambda(\mathcal{I})\}^{-1} \int_{\mathcal{I}} \mathbf{O}(\mathbf{T}(\mathbf{X}(t)), F_{\mathbf{T}(\mathbf{X}(t))})\mathrm{d}t \\
&= \{\lambda(\mathcal{I})\}^{-1} \int_{\mathcal{I}} \mathbf{A}_0 \cdot \mathbf{O}(\mathbf{X}(t), F_{\mathbf{X}(t)})\mathrm{d}t \\
&= \mathbf{A}_0 \cdot \mathbf{MO}(\mathbf{X}, F_{\mathbf{X}}).
\end{aligned}$$

Consequently, we have $\mathbf{MO}(\mathbf{T}(\mathbf{X}_g), F_{\mathbf{T}(\mathbf{X}_g)}) = \mathbf{A}_0 \cdot \mathbf{MO}(\mathbf{X}, F_{\mathbf{X}})$. Similarly,

$$\begin{aligned}
\mathbf{FOM}(\mathbf{T}(\mathbf{X}), F_{\mathbf{T}(\mathbf{X})}) &= \{\lambda(\mathcal{I})\}^{-1} \int_{\mathcal{I}} \mathbf{O}(\mathbf{T}(\mathbf{X}(t)), F_{\mathbf{T}(\mathbf{X}(t))})\mathbf{O}^{\mathrm{T}}(\mathbf{T}(\mathbf{X}(t)), F_{\mathbf{T}(\mathbf{X}(t))})\mathrm{d}t \\
&= \mathbf{A}_0 \left\{ \{\lambda(\mathcal{I})\}^{-1} \int_{\mathcal{I}} \mathbf{O}(\mathbf{X}(t), F_{\mathbf{X}(t)})\mathbf{O}^{\mathrm{T}}(\mathbf{X}(t), F_{\mathbf{X}(t)})\mathrm{d}t \right\} \mathbf{A}_0^{\mathrm{T}} \\
&= \mathbf{A}_0\mathbf{FOM}(\mathbf{X}, F_{\mathbf{X}})\mathbf{A}_0^{\mathrm{T}}.
\end{aligned}$$

This leads to $\mathbf{FOM}\left(\mathbf{T}(\mathbf{X}_g), F_{\mathbf{T}(\mathbf{X}_g)}\right) = \mathbf{A}_0\mathbf{FOM}(\mathbf{X}, F_{\mathbf{X}})\mathbf{A}_0^{\mathrm{T}}$. Finally,

$$\begin{aligned}
\mathbf{VOM}\left(\mathbf{T}(\mathbf{X}_g), F_{\mathbf{T}(\mathbf{X}_g)}\right) &= \mathbf{A}_0 \left\{ \mathbf{FOM}(\mathbf{X}, F_{\mathbf{X}}) - \mathbf{MO}(\mathbf{X}, F_{\mathbf{X}})\mathbf{MO}^{\mathrm{T}}(\mathbf{X}, F_{\mathbf{X}}) \right\} \mathbf{A}_0^{\mathrm{T}} \\
&= \mathbf{A}_0\mathbf{VOM}(\mathbf{X}, F_{\mathbf{X}})\mathbf{A}_0^{\mathrm{T}}
\end{aligned}$$

This completes the proof. $\square$