Statistical Depth Functions

Satvik Saha
Supervised by Dr. Anirvan Chakraborty

17 May, 2024

Department of Mathematics and Statistics,
Indian Institute of Science Education and Research, Kolkata

Outline

1. Depth Functions for Functional Data
2. Outlier Detection for Functional Data
3. Local Depth Functions
4. Concluding Remarks

A *depth function* quantifies how *central* a point $\boldsymbol{x} \in \mathscr{X}$ is with respect to a distribution F.

This induces a *center-outwards* ordering on the space $\mathscr{X}$.

Depth Functions in $\mathbb{R}^d$

We want $D\colon \mathbb{R}^d \times \mathscr{F} \to \mathbb{R}$ to be bounded, non-negative, continuous, and satisfy the following properties.

P1. **Affine invariance:** $D(A\mathbf{x} + b, F_{A\mathbf{x}+b}) = D(\mathbf{x}, F_X)$.

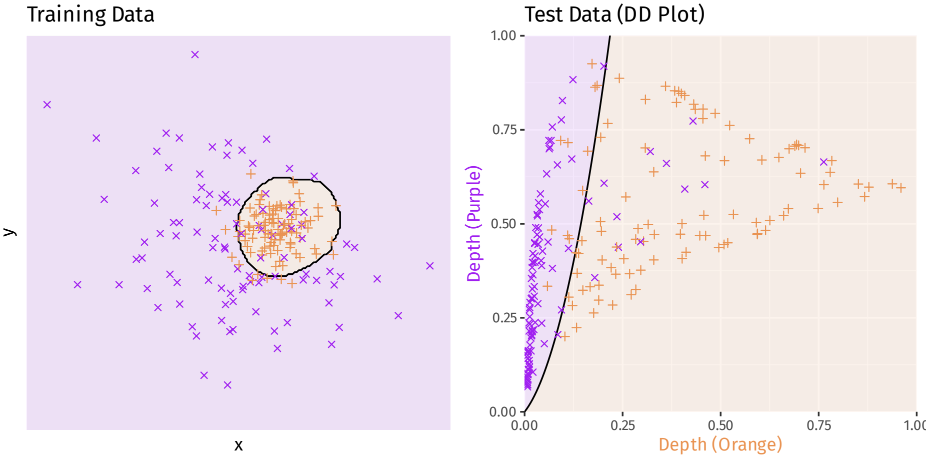
P2. **Maximality at centre:** $D(\theta, F_X) = \sup_{\mathbf{x}\in\mathbb{R}^d} D(\mathbf{x}, F)$.

P3. **Monotonicity along rays:** $D(\mathbf{x}, F) \leq D(\theta + \alpha(\mathbf{x} - \theta), F)$.

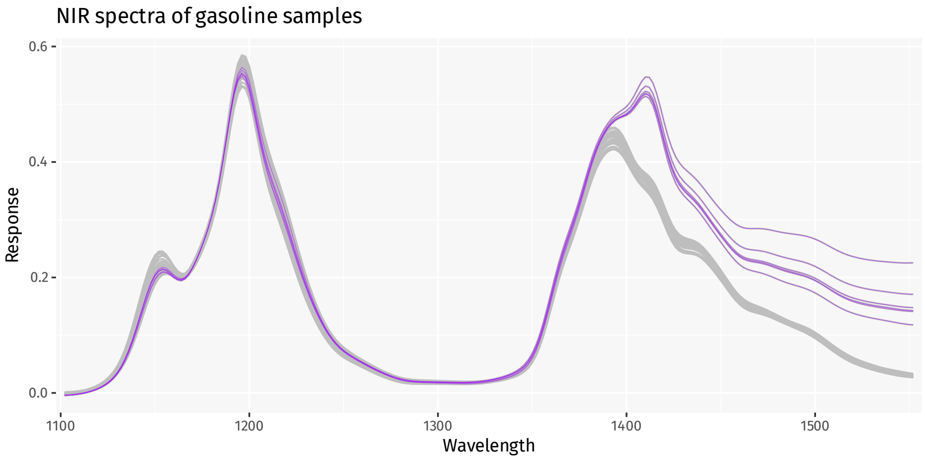
P4. **Vanish at infinity:** $D(\mathbf{x}, F) \to 0$ as $\|\mathbf{x}\| \to \infty$.

Zuo, Y., & Serfling, R. (2000) General notions of statistical depth function

The DD classifier



Depth Functions for Functional Data



NIR spectra of gasoline samples
0.6
0.4
0.2
0.0
Response
1100
1200
1300
1400
1500
Wavelength

Depth Functions in Banach spaces $\mathscr{X}$

Let $\mathscr{X}$ be a class of functions of the form $\boldsymbol{x}\colon [0,1] \to \mathbb{R}^d$, equipped with a norm $\|\cdot\|$. We typically choose $L^2[0,1]$ or $\mathcal{C}[0,1]$.

We want to generalize the Zuo-Serfling properties (P1-4) in this setting, for depth functions $D\colon \mathscr{X} \times \mathscr{F} \to \mathbb{R}$.

Gijbels, I., & Nagy, S. (2017) On a General Definition of Depth for Functional Data

2024-05-17

Depth Functions in Banach spaces $\mathscr{X}$

Let $\mathscr{X}$ be a class of functions of the form $\boldsymbol{x}\colon [0,1] \to \mathbb{R}^d$, equipped with a norm $\|\cdot\|$. We typically choose $L^2[0,1]$ or $\mathcal{C}[0,1]$.

We want to generalize the Zuo-Serfling properties (P1-4) in this setting, for depth functions $D\colon \mathscr{X} \times \mathscr{F} \to \mathbb{R}$.

Gijbels, I., & Nagy, S. (2017) On a General Definition of Depth for Functional Data

Properties $P3$ (Monotonicity along rays) and $P4$ (Vanish at infinity) carry over naturally.

P0. **Non-degeneracy:** $\inf_{x\in\mathcal{X}} D(x,F) < \sup_{x\in\mathcal{X}} D(x,F)$.

The naïve generalization of the halfspace/Tukey depth

$$D_H(x,F) = \inf_{v\in\mathcal{X}^*} P_{X\sim F}(v^*(X) \leq v^*(x)),$$

is degenerate for a wide class of distributions $\mathcal{F}$. For instance, $\mathcal{X} = \mathcal{C}[0,1]$, Gaussian processes with positive definite covariance kernels.

Chakraborty, A., & and Chaudhuri, P. (2014) On data depth in infinite dimensional spaces

Non-degeneracy

P0. **Non-degeneracy:** $\inf_{x \in \mathscr{X}} D(x, F) < \sup_{x \in \mathscr{X}} D(x, F)$.

The naïve generalization of the halfspace/Tukey depth

$$D_H(x, F) = \inf_{v \in \mathscr{X}^*} P_{X \sim F}(v^*(X) \leq v^*(x)),$$

is degenerate for a wide class of distributions $\mathscr{F}$. For instance, $\mathscr{X} = \mathcal{C}[0, 1]$, Gaussian processes with positive definite covariance kernels.

Chakraborty, A., & and Chaudhuri, P. (2014) On data depth in infinite dimensional spaces

This also applies to the functional analogue of the projection depth.

Non-degeneracy

The functional analogue of the spatial depth

$$D_{Sp}(x, F) = 1 - \left\| \mathbb{E}_{X \sim F} \left[\frac{x - X}{\|x - X\|_2} \right] \right\|_2 ,$$

does not suffer from degeneracy.

Chakraborty, A., & and Chaudhuri, P. (2014) The spatial distribution in infinite dimensional spaces and related quantiles and depths

Affine Invariance

P1S. **Scalar-affine invariance:** For $a, b \in \mathbb{R}$ with a non-zero and $\mathbf{x} \in \mathscr{X}$,

$$D(a\mathbf{x} + b, F_{a\mathbf{X}+b}) = D(\mathbf{x}, F_{\mathbf{X}}).$$

P1F. **Function-affine invariance:** For $a, b, x \in \mathscr{X}$, with $ax \in \mathscr{X}$,

$$D(ax + b, F_{aX+b}) = D(x, F_X).$$

Maximality at Center

We say that F_X is symmetric about $\boldsymbol{\theta} \in \mathscr{X}$ if for all $\varphi \in \mathscr{X}^*$, we have $\varphi(X)$ symmetric about $\varphi(\boldsymbol{\theta})$.

P2C. **Maximality at center of central symmetry:** For $F \in \mathscr{X}$ centrally symmetric about $\boldsymbol{\theta} \in \mathscr{X}$, $D(\boldsymbol{\theta}, F) = \sup_{x \in \mathscr{X}} D(x, F)$.

P2H. **Maximality at center of halfspace symmetry:** For $F \in \mathscr{X}$ halfspace symmetric about $\boldsymbol{\theta} \in \mathscr{X}$, $D(\boldsymbol{\theta}, F) = \sup_{x \in \mathscr{X}} D(x, F)$.

The Integrated and Infimal Depths

$$D_{FM}(\mathbf{x}, F_{\mathbf{X}}) = \int_{[0,1]} D(\mathbf{x}(t), F_{\mathbf{X}(t)})\, w(t)\, dt.$$

$$D_{Inf}(\mathbf{x}, F_{\mathbf{X}}) = \inf_{t \in [0,1]} D(\mathbf{x}(t), F_{\mathbf{X}(t)}).$$

Fraiman, R., & and Muniz, G. (2001) Trimmed means for functional data
Mosler, K. (2013) Depth Statistics

The J-th order Integrated and Infimal Depths

$$D^J_{FM}(x, F_X) = \int_{[0,1]^J} D((x(t_1), \ldots, x(t_J))^\top, F_{(X(t_1),\ldots,X(t_J))^\top})\, w(\mathbf{t})\, d\mathbf{t}.$$

$$D^J_{Inf}(x, F_X) = \inf_{\mathbf{t}\in[0,1]^J} D((x(t_1), \ldots, x(t_J))^\top, F_{(X(t_1),\ldots,X(t_J))^\top}).$$

Nagy, S., Gijbels, I., & and Hlubinka, D. (2017) Depth-Based Recognition of Shape Outlying Functions

The J-th order Integrated and Infimal Depths

$$D^J_{FM}(\mathbf{x}, F_X) = \int_{[0,1]^J} D((x(t_1), \dots, x(t_J))^\top, F_{(X(t_1),\dots,X(t_J))^\top})\, w(\mathbf{t})\, d\mathbf{t}.$$

$$D^J_{Inf}(\mathbf{x}, F_X) = \inf_{\mathbf{t} \in [0,1]^J} D((x(t_1), \dots, x(t_J))^\top, F_{(X(t_1),\dots,X(t_J))^\top}).$$

Nagy, S., Gijbels, I., & and Hlubinka, D. (2017) Depth-Based Recognition of Shape Outlying Functions

These J-th order depths carry information about the derivatives of the curves, of orders $0, \dots, J-1$.

The Band Depth

$$D_B^J(\boldsymbol{x}, F) = \sum_{j=2}^{J} P_{\boldsymbol{X}_i \overset{\text{iid}}{\sim} F}(\boldsymbol{x} \in \text{conv}(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_j)).$$

This is the proportion of j-tuples of curves, for $2 \leq j \leq J$, which *completely* envelope $\boldsymbol{x}$.

The band depth becomes degenerate for $\mathscr{X} = \mathcal{C}[0,1]$, Feller processes $\boldsymbol{X}$ (e.g. Brownian motion) with $P(\boldsymbol{X}_0 = 0) = 1$ and each $\boldsymbol{X}_t$ for $t > 1$ non-atomic and symmetric about 0.

López Pintado, S., & Romo, J. (2009) On the concept of depth for functional data

The Modified Band Depth

Define the *enveloping time*

$$\mathrm{ET}(x; x_1, \ldots, x_j) = m_1(\{t \in [0,1] \colon x(t) \in \mathrm{conv}(x_1(t), \ldots, x_j(t))\})$$

The modified band depth is defined as

$$D_{\mathrm{MBD}}(x, F) = \sum_{j=2}^{J} \mathbb{E}_{X_i \overset{\mathrm{iid}}{\sim} F} \left[\mathrm{ET}(x; X_1, \ldots, X_j) \right].$$

The Half-Region Depth

We say that $\boldsymbol{y}$ is in the hypograph (resp. epigraph) of $\boldsymbol{x}$, denoted $\boldsymbol{y} \in H_{\boldsymbol{x}}$ (resp. $E_{\boldsymbol{x}}$), if $\boldsymbol{y}(t) \leq \boldsymbol{x}(t)$ (resp. $\geq$) for all $t \in [0, 1]$.

The half-region depth is defined as

$$D_{HR}(\boldsymbol{x}, F) = \min\left\{P_F(H_{\boldsymbol{x}}),\ P_F(E_{\boldsymbol{x}})\right\}.$$

This suffers from the same degeneracy problems as the band depth.

López Pintado, S., & Romo, J. (2011) A half-region depth for functional data

The Modified Half-Region Depth

Define the Modified Hypograph (MHI) and Epigraph (MEI) Indices as

$$\mathrm{MHI}_F(x) = \mathbb{E}_{X \in F}[m_1\{t \in [0,1] \colon x(t) \geq X(t)\}],$$
$$\mathrm{MEI}_F(x) = \mathbb{E}_{X \in F}[m_1\{t \in [0,1] \colon x(t) \leq X(t)\}].$$

The modified half-region depth is defined as

$$D_{MHR}(x, F) = \min\left\{\mathrm{MHI}_F(x),\ \mathrm{MEI}_F(x)\right\}.$$

Partially Observed Functional Data

Suppose that $X \sim F_X$ is not observed on the entire interval $[0,1]$, but rather on some random compact subinterval $O \sim Q$ (independent of X).

Given a dataset $\mathscr{D} = \{(X_i, O_i)\}_{i=1}^n$ where $(X_i, O_i) \overset{\text{iid}}{\sim} F_X \times Q$, we keep track of the indices observed at time $t \in [0,1]$ as $\mathscr{J}(t) = \{j \colon t \in O_j\}$, as well as their number $q(t) = |\mathscr{J}(t)|$.

The Partially Observed Integrated Functional Depth (POIFD)

We may define a depth function in this setting via

$$D_{POIFD}((X, o), F_X \times Q) = \int_o D(X(t), F_{X(t)}) \, w_o(t) \, dt,$$

where $w_o(t) = q(t) / \int_0 q(t) \, dt$.

Elías, A., Jiménez, R., Paganoni, A. M., & Sangalli, L. M. (2023) Integrated depths for partially observed functional data

The Functional Reconstruction Problem

Given (X, O), can we estimate X on $M = [0, 1] \setminus O$?

We may search for a reconstruction operator $\mathcal{R} \colon L^2(O) \to L^2(M)$ that minimizes the mean integrated prediction squared error loss $\mathbb{E}[\|X_M - \mathcal{R}(X_O)\|^2]$. In this setup, the best predictor is the conditional expectation $\mathbb{E}[X_M \mid X_O]$.

We may also search for a continuous linear reconstruction operator $\mathscr{A}$, by estimating terms of the Karhunen-Loéve expansion of X.

Kneip, A., & Liebl, D. (2020) On the optimal reconstruction of partially observed functional data

The Functional Reconstruction Problem

Another approach is to take a convex linear combination of curves from a suitable curve envelope with indices $\mathcal{I}$.

The enveloping curves $\mathcal{I}$ may be chosen so that (X, O) is as deep as possible inside the curve envelope.

Additionally, we want $\mathcal{I}$ to envelope (X, O) for as long as possible (in the sense of the enveloping time ET), and contain as many near curves (in an appropriately modified norm $\|\cdot\|'$) to (X, O) as possible.

Elías, A., Jiménez, R., & Shang, H. L. (2023) Depth-based reconstruction method for incomplete functional data

Outlier Detection for Functional Data

A Naïve Outlier Detection Scheme

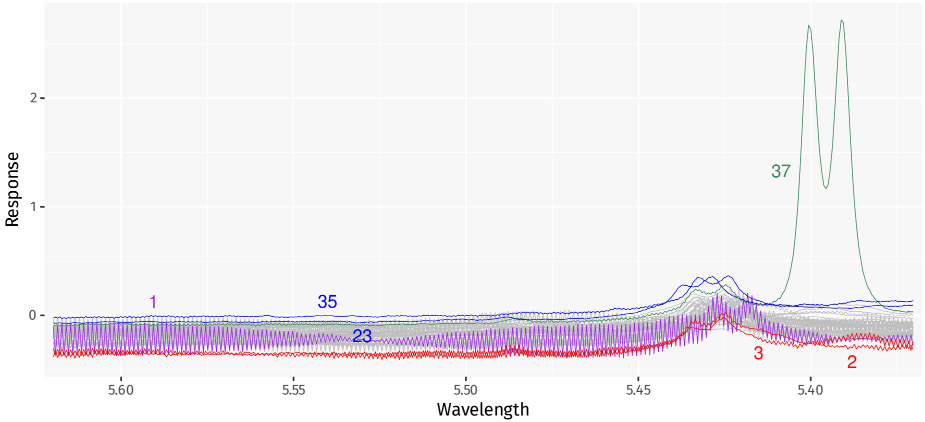
Given data $\mathscr{D} = \{\mathbf{x}_i\}_{i=1}^n$, we may extract ranks $r_i = R(\mathbf{x}_i, \hat{F}_n)$.

For instance, we may choose

$$R(\mathbf{x}, \hat{F}_n) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(D(\mathbf{x}_i, \hat{F}_n) \leq D(\mathbf{x}, \hat{F}_n)).$$

Declare those $\mathbf{x}_i$ with unusually high ranks r_i as outliers, say greater than a cutoff $Q_3 + 1.5\,\mathrm{IQR}$.

NMR spectra of wine samples



Functional Outliers

A curve $x\colon [0,1] \to \mathbb{R}$ may exhibit outlying behaviour within a body of curves in many ways.

- **Isolated outlier:** Significant deviation over a short interval.
- **Persistent outlier:** Deviation over a large/entire interval.
    - Shape
    - Shift
    - Amplitude

Hubert, M., Rousseeuw, P. J., & Segaert, P. (2015) Multivariate functional outlier detection

Functional Outliers

A curve $x\colon [0,1] \to \mathbb{R}$ may exhibit outlying behaviour within a body of curves in many ways.

- **Isolated outlier:** Significant deviation over a short interval.
- **Persistent outlier:** Deviation over a large/entire interval.
  - Shape
  - Shift
  - Amplitude

Hubert, M., Rousseeuw, P. J., & Segaert, P. (2015) Multivariate functional outlier detection

For a shape outlier, the slices $x(t)$ may all seem inconspicuous in the marginals $F_{X(t)}$.

One way of incorporating shape information of a curve $\boldsymbol{x}$ is to bundle it with its derivatives $\boldsymbol{x}^{(j)}$.

$$\int_{[0,1]} D((\boldsymbol{x}^{(0)}(t), \ldots, \boldsymbol{x}^{(J)}(t))^\top, F_{(\boldsymbol{X}^{(0)}(t), \ldots, \boldsymbol{X}^{(J)}(t))^\top})\, w(t)\, dt.$$

Shape Outliers and Derivatives

One way of incorporating shape information of a curve $\boldsymbol{x}$ is to bundle it with its derivatives $\boldsymbol{x}^{(j)}$.

$$\int_{[0,1]} D((x^{(0)}(t), \ldots, x^{(J)}(t))^\top, F_{(X^{(0)}(t),\ldots,X^{(J)}(t))^\top})\, w(t)\, dt.$$

This suffers from errors in approximating derivatives, and the assumption of differentiability in the first place.

Shape Outliers and the J-th order Integrated Depth

We say that a curve $\boldsymbol{x}$ is a J-th order outlier with respect to $F_{\boldsymbol{X}}$ if there exists $\boldsymbol{t} \in [0,1]^J$ such that the vector $(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_J))^\top$ is outlying with respect to $F_{(\boldsymbol{X}(t_1),\ldots,\boldsymbol{X}(t_J))^\top}$.

$$D^J_{FM}(\boldsymbol{x}, F_{\boldsymbol{X}}) = \int_{[0,1]^J} D((\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_J))^\top, F_{(\boldsymbol{X}(t_1),\ldots,\boldsymbol{X}(t_J))^\top})\, w(\boldsymbol{t})\, d\boldsymbol{t}.$$

Nagy, S., Gijbels, I., & and Hlubinka, D. (2017) Depth-Based Recognition of Shape Outlying Functions

Shape Outliers and the J-th order Integrated Depth

We say that a curve $\boldsymbol{x}$ is a J-th order outlier with respect to F_X if there exists $\boldsymbol{t} \in [0,1]^J$ such that the vector $(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_J))^\top$ is outlying with respect to $F_{(X(t_1),\ldots,X(t_J))^\top}$.

$$D^J_{FM}(\boldsymbol{x}, F_X) = \int_{[0,1]^J} D((\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_J))^\top, F_{(X(t_1),\ldots,X(t_J))^\top})\, w(\boldsymbol{t})\, d\boldsymbol{t}.$$

Nagy, S., Gijbels, I., & and Hlubinka, D. (2017) Depth-Based Recognition of Shape Outlying Functions

- This process looks at points of the form $(\boldsymbol{x}(t), \boldsymbol{x}(t+h), \ldots)$, thus encoding information about the derivatives.
- One may choose the weight function $w(\cdot)$ to put emphasis on the diagonal.

The Centrality-Stability Scheme

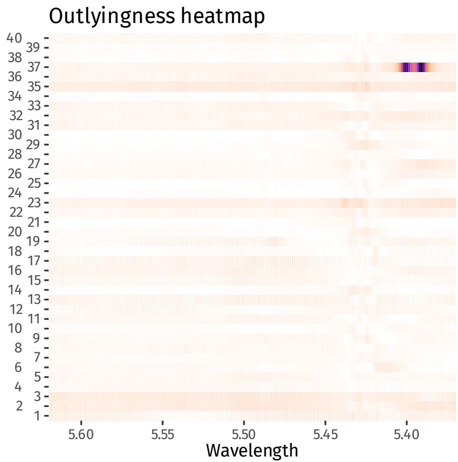
Consider an outlyingness function $O(\mathbf{x}(t))$ which measures the outlyingness of $\mathbf{x}(t)$ with respect to $F_{X(t)}$. For instance, we may choose

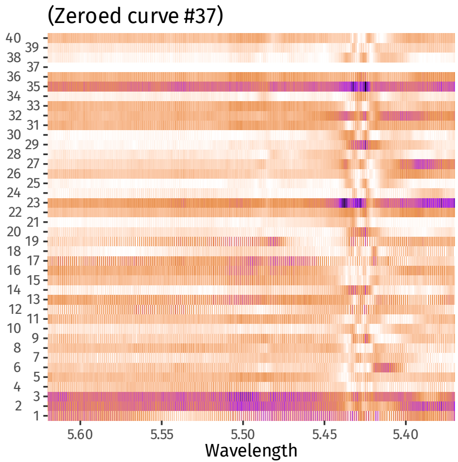
$$O(\mathbf{x}(t)) = \frac{\mathbf{x}(t) - \mathrm{med}(X(t))}{\mathrm{MAD}(X(t))}.$$

Then, we may define a depth function

$$D(\mathbf{x}, F_X) = \int_{[0,1]} (1 + |O(\mathbf{x}(t))|)^{-1}\, dt.$$

Hubert, M., Rousseeuw, P. J., & Segaert, P. (2015) Multivariate functional outlier detection


Outlyingness heatmap
Wavelength
5.60
5.55
5.50
5.45
5.40


(Zeroed curve #37)
Wavelength
5.60
5.55
5.50
5.45
5.40

Define

$$\widetilde{MO}(x, F_X) = \int_{[0,1]} |O(x(t))| \, dt$$

Then, Cauchy-Schwarz gives

$$D(x, F_X) \cdot (1 + \widetilde{MO}(x, F_X)) \geq 1,$$

with equality when $O(x(\cdot))$ remains constant over time.

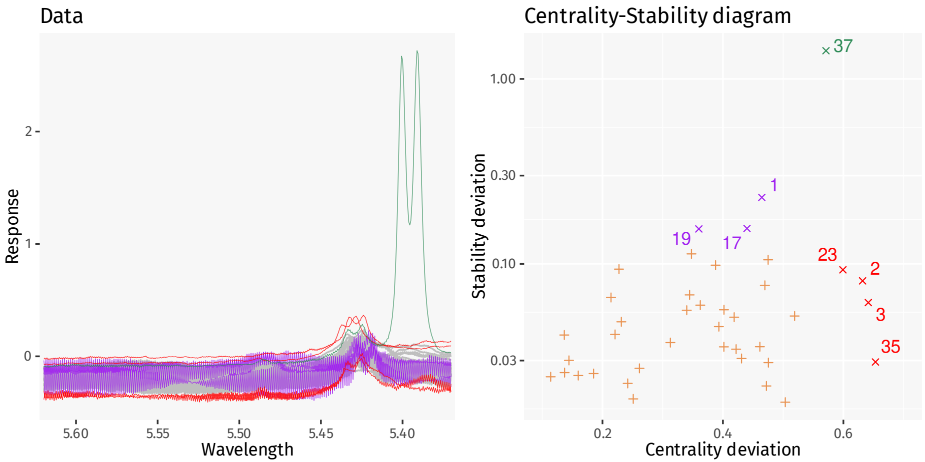
The Centrality-Stability scheme

Any sudden deviation in outlyingness will be detected by the *stability deviation*

$$\Delta S = (1 + \widetilde{MO}(\mathbf{x}, F)) - \frac{1}{D(\mathbf{x}, F)}.$$

The *centrality deviation* is measured as

$$\Delta C = 1 - D(\mathbf{x}, F).$$

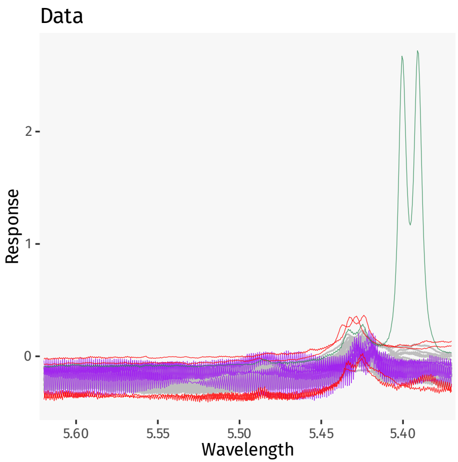

Data
Response
2
1
0
5.60
5.55
5.50
5.45
5.40
Wavelength
Centrality-Stability diagram
Stability deviation
1.00
0.30
0.10
0.03
0.2
0.4
0.6
Centrality deviation
37
1
19
17
23
2
3
35

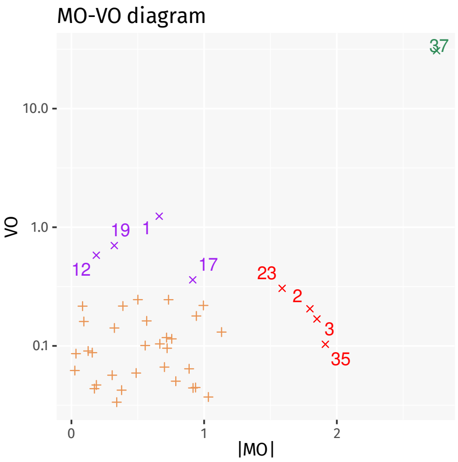
The MO-VO Scheme

We may measure the variability in outlyingness over time more simply via

$$\mathrm{MO}(\mathbf{x}, F) = \int_{[0,1]} O(\mathbf{x}(t))\, dt,$$
$$\mathrm{VO}(\mathbf{x}, F) = \int_{[0,1]} \|O(\mathbf{x}(t)) - \mathbf{MO}(\mathbf{x}, F)\|^2\, dt.$$

Dai, W., & Genton, M. G. (2018) An outlyingness matrix for multivariate functional data classification


Data
Response
2
1
0
5.60
5.55
5.50
5.45
5.40
Wavelength


MO-VO diagram
VO
10.0
1.0
0.1
0
1
2
|MO|
37
19
1
12
17
23
2
3
35

Given a dataset of curves $\mathscr{D} = \{\boldsymbol{x}_i\}_{i=1}^{n}$, the distances

$$d_i = a_0 + a_1 \operatorname{MEI}(\boldsymbol{x}_i) + a_2 n^2 \operatorname{MEI}(\boldsymbol{x}_i)^2 - \operatorname{MBD}(\boldsymbol{x}_i),$$

where $a_0 = a_2 = -2/n(n+1)$, $a_1 = 2(n+1)/(n-1)$, are indicative of shape outlyingness.

Thus, one may declare $\boldsymbol{x}_i$ as an outlier if d_i exceeds a cutoff such as $Q_3 + 1.5\,\mathrm{IQR}$.

Arribas-Gil, A., & Romo, J. (2014). Shape outlier detection and visualization for functional data: the outliergram

2024-05-17

The numbers d_i are always positive!

The Outliergram

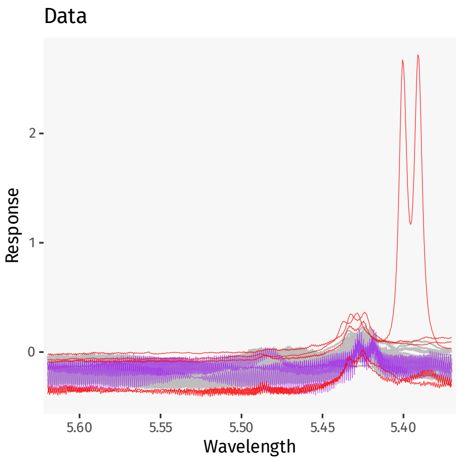
Given a dataset of curves $\mathcal{D} = \{x_i\}_{i=1}^n$, the distances

$$d_i = a_0 + a_1 \operatorname{MEI}(x_i) + a_2 n^2 \operatorname{MEI}(x_i)^2 - \operatorname{MBD}(x_i),$$

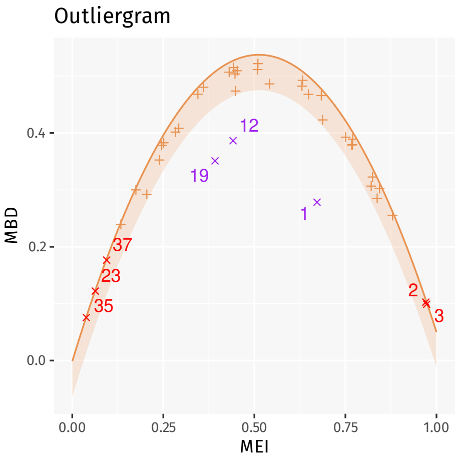
where $a_0 = a_2 = -2/n(n+1)$, $a_1 = 2(n+1)/(n-1)$, are indicative of shape outlyingness.

Thus, one may declare x_i as an outlier if d_i exceeds a cutoff such as $Q_3 + 1.5\,\mathrm{IQR}$.

Arribas-Gil, A., & Romo, J. (2014). Shape outlier detection and visualization for functional data: the outliergram



Data
Response
2
1
0
5.60
5.55
5.50
5.45
5.40
Wavelength



Outliergram
MBD
0.4
0.2
0.0
12
19
1
37
23
35
2
3
0.00
0.25
0.50
0.75
1.00
MEI

Local Depth Functions

We say that a distribution F is elliptical if it admits a density of the form

$$f_X(x) = c|\Sigma|^{-1/2} h\left((x - \boldsymbol{\mu})^\top \Sigma^{-1} (x - \boldsymbol{\mu})\right)$$

for some strictly decreasing function h. Write $F \in \mathrm{Ell}(h; \boldsymbol{\mu}, \Sigma)$.

An affine invariant depth function continuous in $\boldsymbol{x}$ uniquely determines F within $\mathrm{Ell}(h; \cdot, \cdot)$. The depth and density contours coincide.

Elliptical Distributions

We say that a distribution F is elliptical if it admits a density of the form

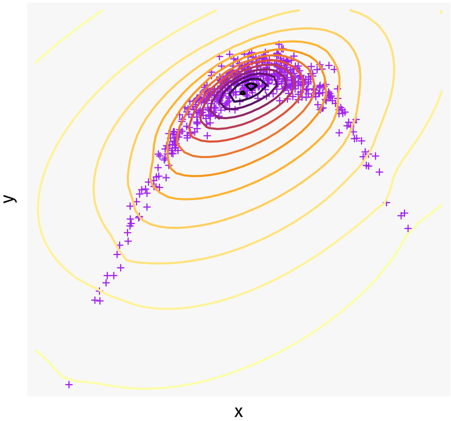
$$f_X(\boldsymbol{x}) = c|\Sigma|^{-1/2} h\left((\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

for some strictly decreasing function h. Write $F \in \mathrm{Ell}(h; \boldsymbol{\mu}, \Sigma)$.

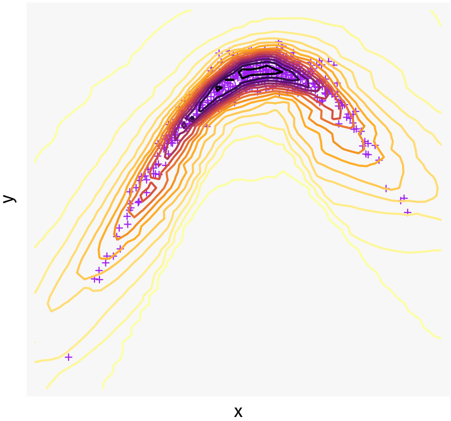
An affine invariant depth function continuous in $\boldsymbol{x}$ uniquely determines F within $\mathrm{Ell}(h; \cdot, \cdot)$. The depth and density contours coincide.

- The whitened random variable $Z = \Sigma^{-1/2}(X - \boldsymbol{\mu})$ has density $f_Z(\boldsymbol{z}) \propto h(\|\boldsymbol{z}\|^2)$.
- The halfspace, simplicial, projection depths satisfy this property.
- In general, depths such as the halfspace depth always produce convex central regions.

Spatial depth



Local spatial depth, β = 0.2



Local Depth Neighbourhoods

Given $\boldsymbol{x} \in \mathscr{X}$, we may symmetrize $F_{\boldsymbol{X}}$ as

$$F_{\boldsymbol{X}}^{\boldsymbol{x}} = \frac{1}{2}F_{\boldsymbol{X}} + \frac{1}{2}F_{2\boldsymbol{x}-\boldsymbol{X}}.$$

The probability-β depth-based neighbourhood of $\boldsymbol{x}$ in $F_{\boldsymbol{X}}$ is simply the β-th central region of $F_{\boldsymbol{X}}^{\boldsymbol{x}}$. This is denoted by $N_{\beta}^{\boldsymbol{x}}(F_{\boldsymbol{X}})$.

Paindaveine, D., & Van Bever, G. (2013) From depth to local depth: A focus on centrality

Local Depth

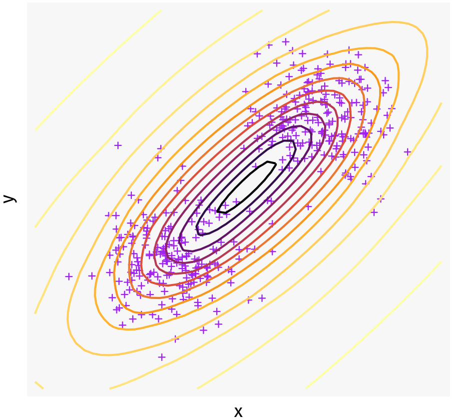
Let F_β^x denote the distribution F_X conditioned on $N_\beta^x(F_X)$.

The local depth function at locality level $\beta \in (0, 1]$ is defined as

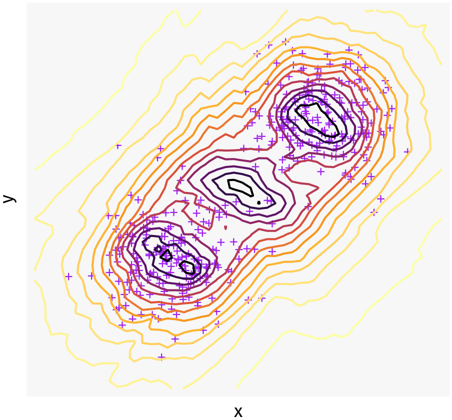
$$LD(x, F_X) = D(x, F_\beta^x).$$

When $\beta = 1$, we have $LD_1 = D$.

Spatial depth



Local spatial depth, β = 0.2



Local Depth based Regression

Let $\widetilde{F}_\beta^{\boldsymbol{x}}$ denote the distribution $F_{\boldsymbol{X}}^{\boldsymbol{x}}$ conditioned on $N_\beta^{\boldsymbol{x}}(F_{\boldsymbol{X}})$. Note that this is angularly symmetric about $\boldsymbol{x}$.

Given $\boldsymbol{x} \in \mathscr{X}$, we may define a local depth kernel, centered at $\boldsymbol{x}$, via

$$K_\beta^{\boldsymbol{x}} \colon N_\beta^{\boldsymbol{x}}(F_{\boldsymbol{X}}) \to \mathbb{R}, \qquad \boldsymbol{z} \mapsto D(\boldsymbol{z}, \widetilde{F}_\beta^{\boldsymbol{x}}).$$

Extend this to $\mathscr{X}$ by setting $K_\beta^{\boldsymbol{x}}(\cdot) = 0$ outside $N_\beta^{\boldsymbol{x}}(F_{\boldsymbol{X}})$.

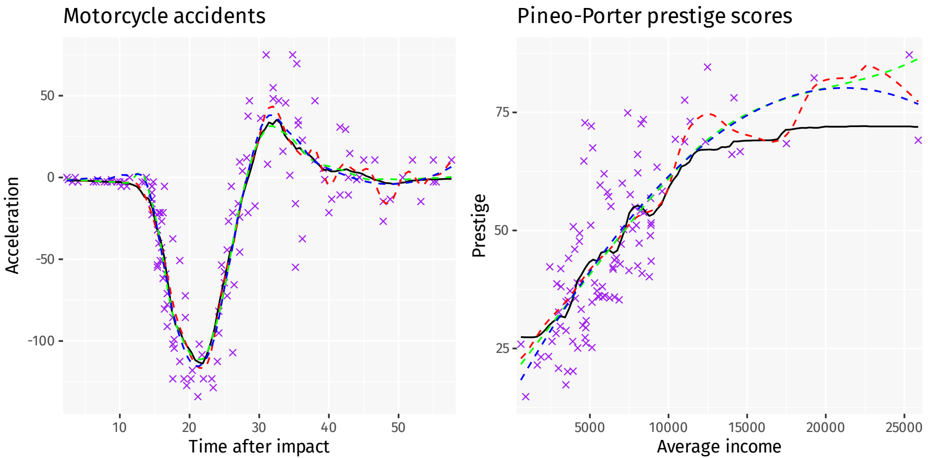
Local Depth based Regression

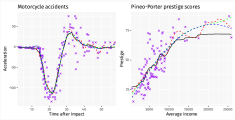
We propose a linear estimator of the form

$$\hat{y}_\beta(x) = \sum_i w_i(x) y_i, \qquad w_i(x) = \frac{K_\beta^x(x_i)}{\sum_j K_\beta^x(x_j)}.$$

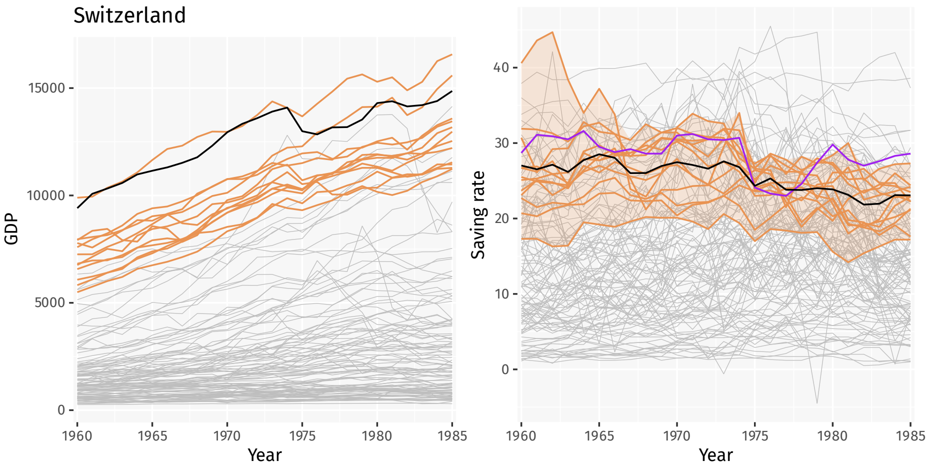
This may be interpreted as a weighted KNN estimator, or a variable bandwidth kernel estimator.

We only have one tuning parameter $\beta \in (0, 1]$.

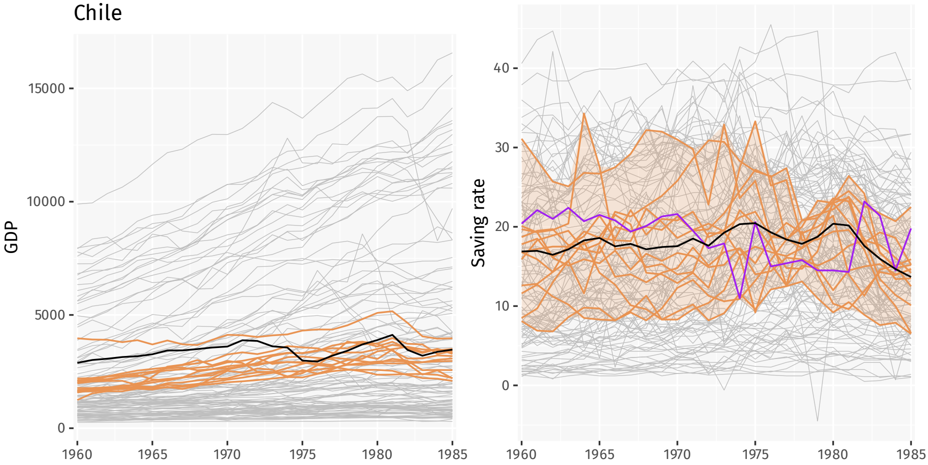

Motorcycle accidents
Acceleration
Time after impact
50
0
-50
-100
10
20
30
40
50
Pineo-Porter prestige scores
Prestige
Average income
75
50
25
5000
10000
15000
20000
25000



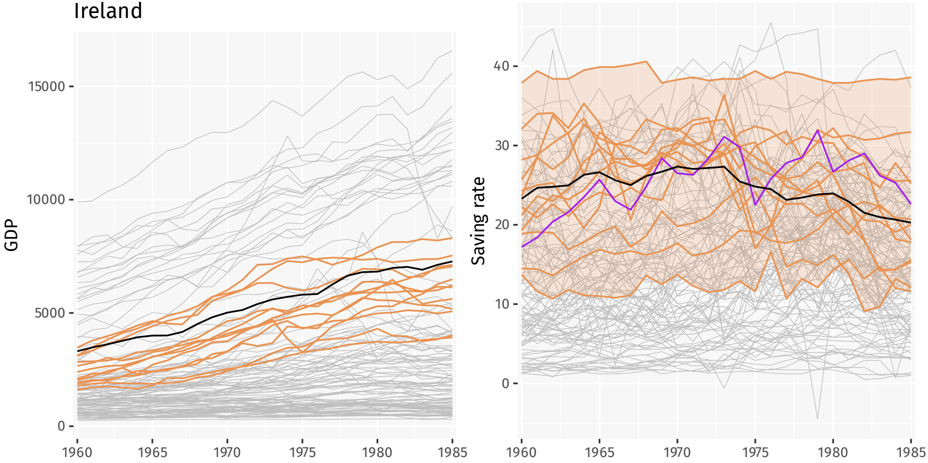
The methods used are local depth based regression (black), Nadaraya-Watson kernel (red), local linear (green) and quadratic (blue).



Switzerland
15000
10000
5000
0
GDP
1960
1965
1970
1975
1980
1985
Year
40
30
20
10
0
Saving rate
1960
1965
1970
1975
1980
1985
Year

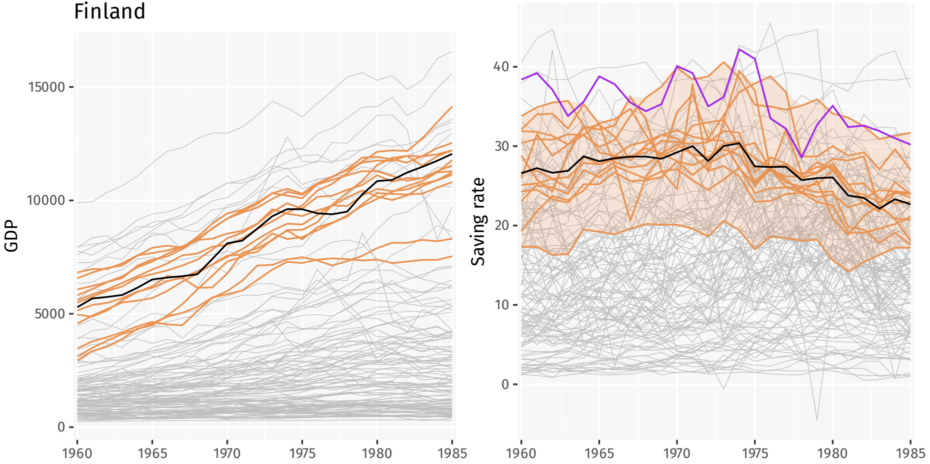

Chile
GDP
15000
10000
5000
0
1960
1965
1970
1975
1980
1985
Saving rate
40
30
20
10
0
1960
1965
1970
1975
1980
1985

Ireland



GDP
15000
10000
5000
0
1960
1965
1970
1975
1980
1985
Saving rate
40
30
20
10
0
1960
1965
1970
1975
1980
1985

Finland


GDP
15000
10000
5000
0
1960
1965
1970
1975
1980
1985
Saving rate
40
30
20
10
0
1960
1965
1970
1975
1980
1985

Concluding Remarks

- Formulation and Properties of Depth Functions
  - Multivariate Data
  - Functional Data
- Extensions of Depth Functions
  - Partially Observed Functional Data
  - Local Depth
- Applications
  - Exploratory Data Analysis
  - Testing
  - Classification
  - Clustering
  - Outlier Detection
  - Data Reconstruction
  - Regression ?

References i

[1] Ana Arribas-Gil and Juan Romo.
Shape outlier detection and visualization for functional data: the outliergram.
Biostatistics, 15(4):603–619, 03 2014.

[2] Anirvan Chakraborty and Probal Chaudhuri.
On data depth in infinite dimensional spaces.
Annals of the Institute of Statistical Mathematics, 66(2):303–324, 2014.

[3] Anirvan Chakraborty and Probal Chaudhuri.
The spatial distribution in infinite dimensional spaces and related quantiles and depths.
The Annals of Statistics, 42(3):1203–1231, 2014.

[4] Wenlin Dai and Marc G. Genton.
An outlyingness matrix for multivariate functional data classification.
Statistica Sinica, 28(4):2435–2454, 2018.

[5] Antonio Elías, Raúl Jiménez, and Han Lin Shang.
Depth-based reconstruction method for incomplete functional data.
Computational Statistics, 38(3):1507–1535, 2023.

[6] Antonio Elías, Raúl Jiménez, Anna M. Paganoni, and Laura M. Sangalli.
Integrated depths for partially observed functional data.
Journal of Computational and Graphical Statistics, 32(2):341–352, 2023.

[7] Ricardo Fraiman and Graciela Muniz.
Trimmed means for functional data.
Test, 10(2):419–440, 2001.

[8] Irène Gijbels and Stanislav Nagy.
On a General Definition of Depth for Functional Data.
Statistical Science, 32(4):630 – 639, 2017.

[9] Mia Hubert, Peter J. Rousseeuw, and Pieter Segaert.
Multivariate functional outlier detection.
Statistical Methods & Applications, 24(2):177–202, 2015.

[10] Alois Kneip and Dominik Liebl.
On the optimal reconstruction of partially observed functional data.
The Annals of Statistics, 28(3):1692–1717, 2020.

[11] David Kraus.
Components and completion of partially observed functional data.
Journal of the Royal Statistical Society. Series B (Statistical Methodology), 77(4):777–801, 2015.

[12] Sara López-Pintado and Juan Romo.
On the concept of depth for functional data.
Journal of the American Statistical Association, 104(486):718–734, 2009.

[13] Sara López-Pintado and Juan Romo.
A half-region depth for functional data.
Computational Statistics & Data Analysis, 55(4):1679–1695, 2011.

[14] Karl Mosler.
Depth Statistics, pages 17–34.
Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[15] Stanislav Nagy, Irene Gijbels, and Daniel Hlubinka.
Depth-Based Recognition of Shape Outlying Functions.
Journal of Computational and Graphical Statistics, 26(4):883–893, 2017.

[16] Davy Paindaveine and Germain Van Bever.
From depth to local depth: A focus on centrality.
Journal of the American Statistical Association, 108(503):1105–1119, 2013.

[17] Yijun Zuo and Robert Serfling.
General notions of statistical depth function.
The Annals of Statistics, 28(2):461–482, 2000.