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

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 $\mathscr{X}$.

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

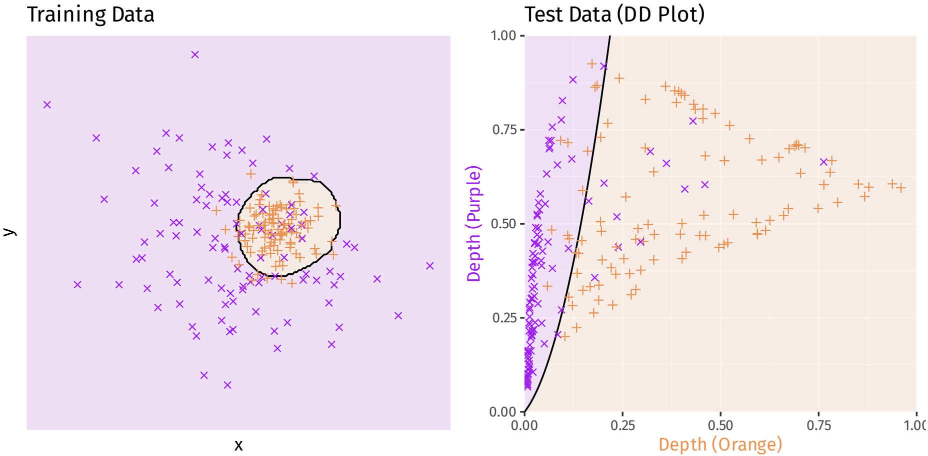
P2. **Maximality at centre:** $D(\theta, F_X) = \sup_{\mathbf{x}\in\mathbb{R}^d} D(\mathbf{x}, F)$.

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

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

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.

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

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

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

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

The J-th order Integrated and Infimal Depths

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

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

The J-th order Integrated and Infimal Depths

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

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

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

The Band Depth

$$D_B^J(x, F) = \sum_{j=2}^{J} P_{X_i \overset{\text{iid}}{\sim} F}(x \in \operatorname{conv}(X_1, \ldots, X_j)).$$

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

The band depth becomes degenerate for $\mathscr{X} = \mathcal{C}[0,1]$, Feller processes X (e.g. Brownian motion) with $P(X_0 = 0) = 1$ and each X_t for $t > 1$ non-atomic and symmetric about 0.

The Modified Band Depth

Define the *enveloping time*

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

The modified band depth is defined as

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

The Half-Region Depth

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

The half-region depth is defined as

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

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

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

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

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

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.

The functional reconstruction problem

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

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

Additionally, we want $\mathscr{I}$ 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.

Outlier detection for Functional Data

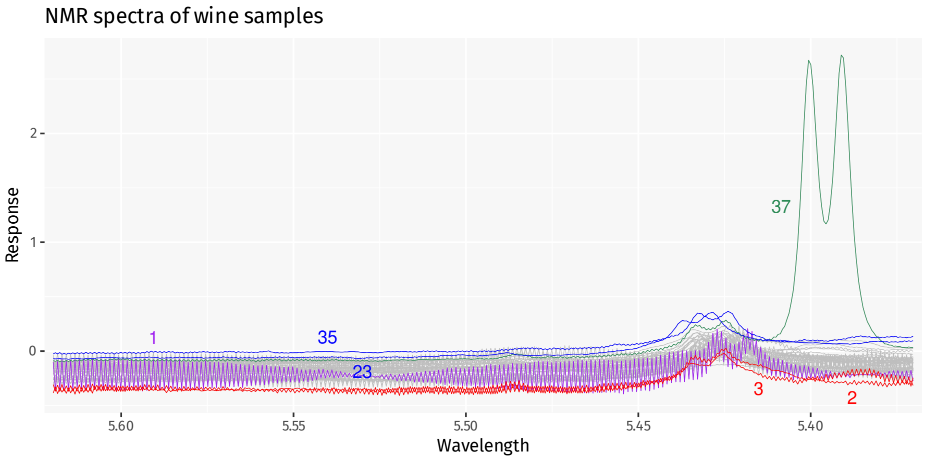
A naïve outlier detection scheme

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
Response
Wavelength
2
1
0
5.60
5.55
5.50
5.45
5.40
37
1
35
23
3
2

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

A curve $x : [0, 1] \to \mathbb{R}$ 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

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

Shape outliers and derivatives

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

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

Shape outliers and derivatives

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

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

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 $F_{(\boldsymbol{X}(t_1),\ldots,\boldsymbol{X}(t_J))^\top}$.

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

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 $(\boldsymbol{x}(t_1), \dots, \boldsymbol{x}(t_J))^\top$ is outlying with respect to $F_{(\boldsymbol{X}(t_1),\dots,\boldsymbol{X}(t_J))^\top}$.

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

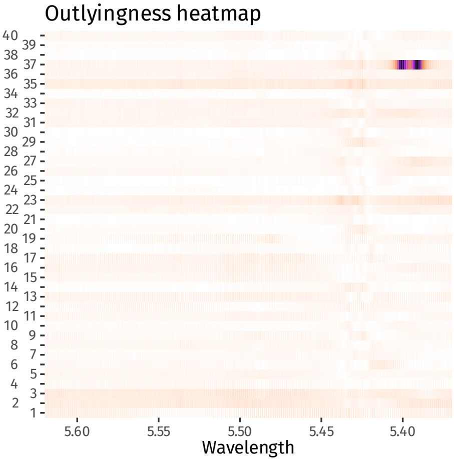
- This process looks at points of the form $(\boldsymbol{x}(t), \boldsymbol{x}(t+h), \dots)$, thus encoding information about the derivatives.
- One may choose the weight function $w(\cdot)$ to put emphasis on the diagonal.

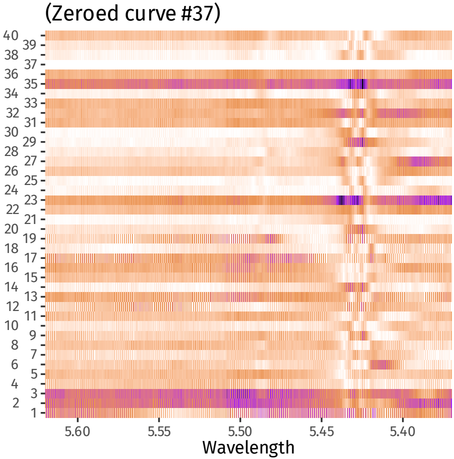
Consider an outlyingness function $O(\mathbf{x}(t))$ which measures the outlyingness of $\mathbf{x}(t)$ with respect to $F_{X(t)}$. For instance, we may choose

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

Then, we may define a depth function

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


Outlyingness heatmap
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
5.60
5.55
5.50
5.45
5.40
Wavelength


(Zeroed curve #37)
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
5.60
5.55
5.50
5.45
5.40
Wavelength

Define

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

Then, Cauchy-Schwarz gives

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

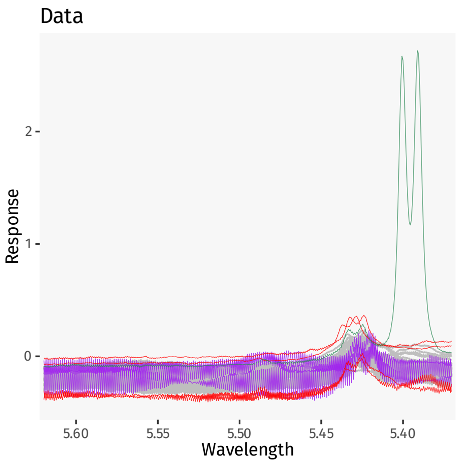
with equality when $O(x(\cdot))$ remains constant over time.

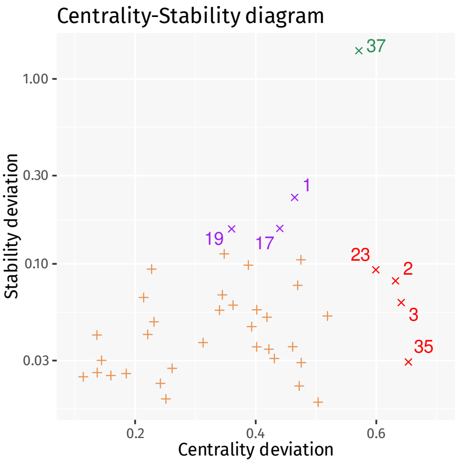
Any sudden deviation in outlyingness will be detected by the *stability deviation*

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

The *centrality deviation* is measured as

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

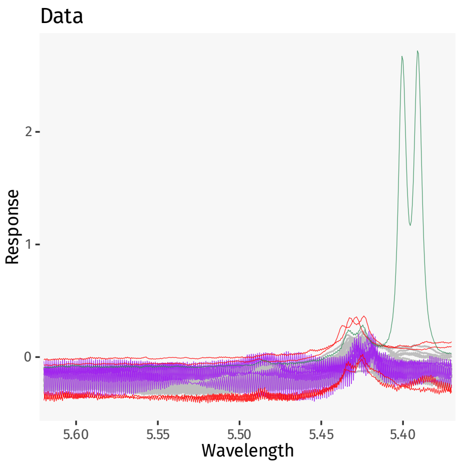

Data
Response
2
1
0
5.60
5.55
5.50
5.45
5.40
Wavelength

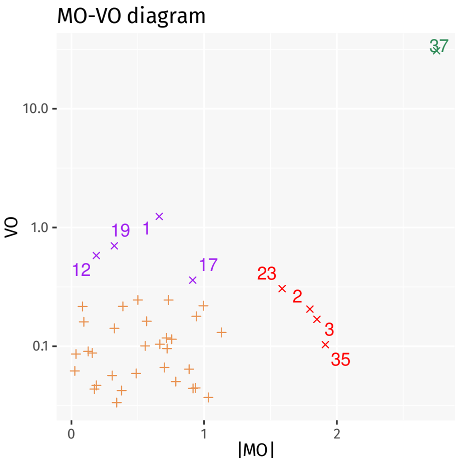

Centrality-Stability diagram
Stability deviation
1.00
0.30
0.10
0.03
0.2
0.4
0.6
Centrality deviation
37
1
19
17
23
2
3
35

The MO-VO scheme

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

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


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

The Outliergram

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

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

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 $\mathbf{x}_i$ as an outlier if d_i exceeds a cutoff such as $Q_3 + 1.5 \operatorname{IQR}$.

The numbers d_i are always positive!

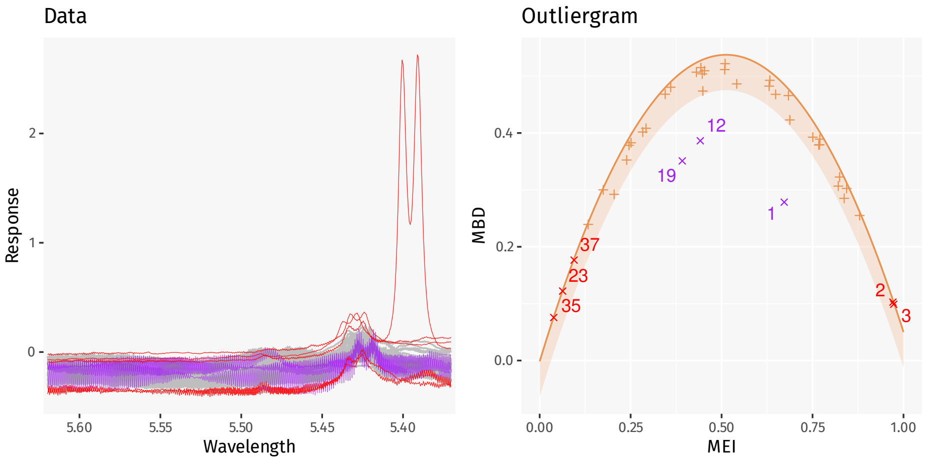
The Outliergram

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

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

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 $\mathbf{x}_i$ as an outlier if d_i exceeds a cutoff such as $Q_3 + 1.5 \operatorname{IQR}$.


Data
Response
Wavelength
2
1
0
5.60
5.55
5.50
5.45
5.40
Outliergram
MBD
MEI
0.4
0.2
0.0
0.00
0.25
0.50
0.75
1.00
12
19
1
37
23
35
2
3

Local Depth Functions

Elliptical distributions

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 x uniquely determines F within $\mathrm{Ell}(h; \cdot, \cdot)$. The depth and density contours coincide.

Elliptical distributions

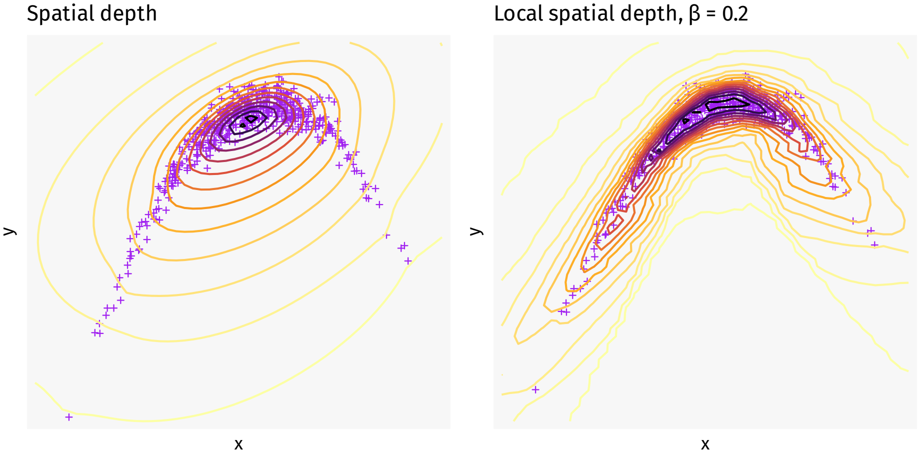
We say that a distribution F is elliptical if it admits a density of the form

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

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 $\boldsymbol{Z} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu})$ 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
y
x
Local spatial depth, β = 0.2
y
x

Local Depth neighbourhoods

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

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

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

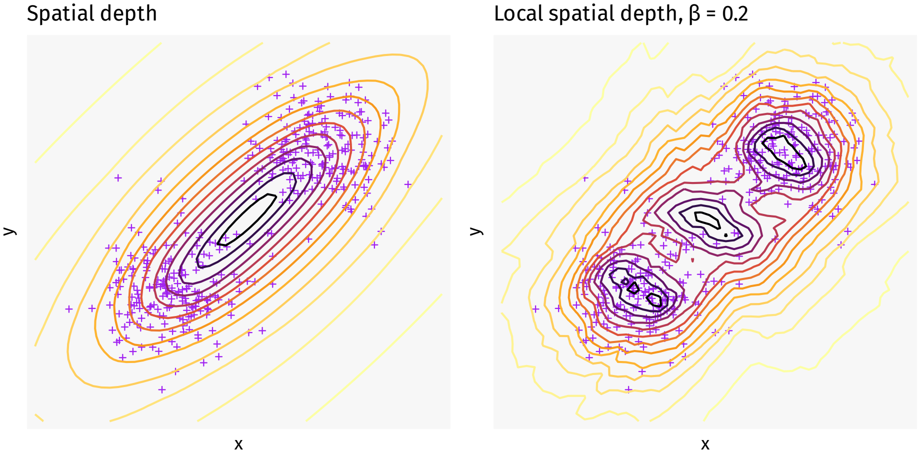
Local Depth

Let $F_\beta^{\mathbf{x}}$ denote the distribution $F_{\mathbf{X}}$ conditioned on $N_\beta^{\mathbf{x}}(F_{\mathbf{X}})$.

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

$$LD(\mathbf{x}, F_{\mathbf{X}}) = D(\mathbf{x}, F_\beta^{\mathbf{x}}).$$

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


Spatial depth
y
x
Local spatial depth, β = 0.2
y
x

Local Depth based Regression

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

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

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

Extend this to $\mathscr{X}$ by setting $K^{\mathbf{x}}_{\beta}(\cdot) = 0$ outside $N^{\mathbf{x}}_{\beta}(F_{\mathbf{X}})$.

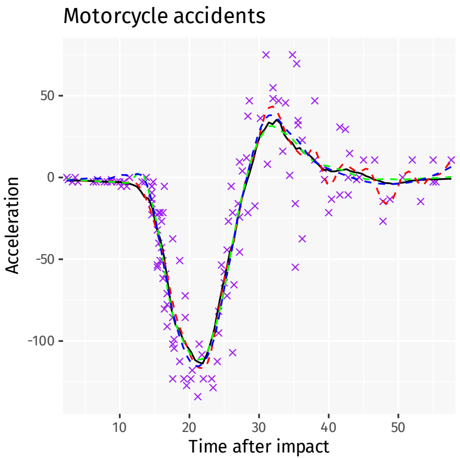
Local Depth based Regression

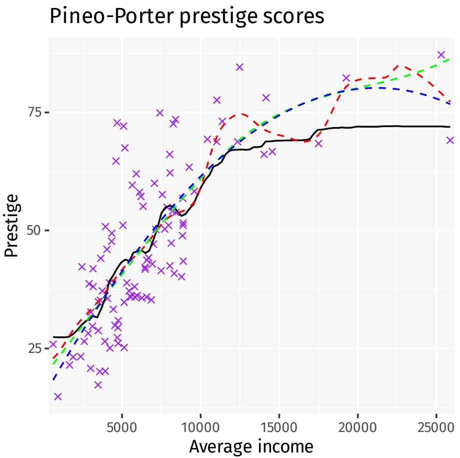
We propose a linear estimator of the form

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

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
50
0
-50
-100
10
20
30
40
50
Time after impact


Pineo-Porter prestige scores
Prestige
75
50
25
5000
10000
15000
20000
25000
Average income

References i

[1] Claudio Agostinelli and Mario Romanazzi.
Local depth.
Journal of Statistical Planning and Inference, 141(2):817–830, 2011.

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

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

[4] 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.

[5] Shojaeddin Chenouri and Christopher G. Small.
A nonparametric multivariate multisample test based on data depth.
Electronic Journal of Statistics, 6(none):760 – 782, 2012.

[6] Victor Chernozhukov, Alfred Galichon, Marc Hallin, and Marc Henry.
Monge–Kantorovich depth, quantiles, ranks and signs.
The Annals of Statistics, 45(1):223 – 256, 2017.

[7] Juan Antonio Cuesta-Albertos, M. Febrero-Bande, and M. Oviedo de la Fuente.
The DD^G-classifier in the functional setting.
TEST, 26(1):119–142, 2017.

[8] Juan Antonio Cuesta-Albertos and Alicia Nieto-Reyes.
The random Tukey depth.
Computational Statistics & Data Analysis, 52(11):4979–4988, 2008.

[9] Juan Antonio Cuesta-Albertos and Alicia Nieto-Reyes.
The Tukey and the random Tukey depths characterize discrete distributions.
Journal of Multivariate Analysis, 99(10):2304–2311, 2008.

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

[11] David L. Donoho and Miriam Gasko.
Breakdown Properties of Location Estimates Based on Halfspace Depth and Projected Outlyingness.
The Annals of Statistics, 20(4):1803 – 1827, 1992.

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

[13] 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.

[14] Antonio Elías, Raúl Jiménez, and Han Lin Shang.
On projection methods for functional time series forecasting.
Journal of Multivariate Analysis, 189:104890, 2022.

[15] Ricardo Fraiman, Regina Y. Liu, and Jean Meloche.
Multivariate Density Estimation by Probing Depth.
Lecture Notes-Monograph Series, 31:415–430, 1997.

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

[17] Yonghong Gao.
Data depth based on spatial rank.
Statistics & Probability Letters, 65(3):217–225, 2003.

[18] Anil K. Ghosh and Probal Chaudhuri.
On maximum depth and related classifiers.
Scandinavian Journal of Statistics, 32(2):327–350, 2005.

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

[20] Xuming He and Gang Wang.
Convergence of Depth Contours for Multivariate Datasets.
The Annals of Statistics, 25(2):495–504, 1997.

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

[22] Rebecka Jörnsten.
Clustering and classification based on the L_1 data depth.
Journal of Multivariate Analysis, 90(1):67–89, 2004.
Special Issue on Multivariate Methods in Genomic Data Analysis.

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

[24] V. I. Koltchinskii.
M-Estimation, Convexity and Quantiles.
The Annals of Statistics, 25(2):435–477, 1997.

[25] Linglong Kong and Yijun Zuo.
Smooth depth contours characterize the underlying distribution.
Journal of Multivariate Analysis, 101(9):2222–2226, 2010.

[26] Gleb Alekseevich Koshevoy.
The Tukey Depth Characterizes the Atomic Measure.
Journal of Multivariate Analysis, 83(2):360–364, 2002.

[27] 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.

[28] Jun Li, Juan Antonio Cuesta-Albertos, and Regina Y. Liu.
DD-Classifier: Nonparametric Classification Procedure Based on DD-Plot.
Journal of the American Statistical Association, 107(498):737–753, 2012.

[29] Regina Y. Liu.
On a Notion of Data Depth Based on Random Simplices.
The Annals of Statistics, 18(1):405 – 414, 1990.

[30] Regina Y. Liu, Jesse M. Parelius, and Kesar Singh.
Multivariate analysis by data depth: descriptive statistics, graphics and inference.
The Annals of Statistics, 27(3):783 – 858, 1999.

[31] Regina Y. Liu and Kesar Singh.
A Quality Index Based on Data Depth and Multivariate Rank Tests.
Journal of the American Statistical Association, 88:252–260, 1993.

[32] 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.

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

[34] Ivan Mizera and Milos Volauf.
Continuity of Halfspace Depth Contours and Maximum Depth Estimators: Diagnostics of Depth-Related Methods.
Journal of Multivariate Analysis, 83(2):365–388, 2002.

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

[36] Karl Mosler and Pavlo Mozharovskyi.
Choosing Among Notions of Multivariate Depth Statistics.
Statistical Science, 37(3):348 – 368, 2022.

[37] Stanislav Nagy.
Monotonicity properties of spatial depth.
Statistics & Probability Letters, 129:373–378, 2017.

[38] Stanislav Nagy.
The halfspace depth characterization problem.
In Michele La Rocca, Brunero Liseo, and Luigi Salmaso, editors, *Nonparametric Statistics*, pages 379–389, Cham, 2020. Springer International Publishing.

[39] Stanislav Nagy.
Halfspace depth does not characterize probability distributions.
Statistical Papers, 62(3):1135–1139, 2021.

[40] 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.

[41] Alicia Nieto-Reyes and Heather Battey.
A Topologically Valid Definition of Depth for Functional Data.
Statistical Science, 31(1):61 – 79, 2016.

[42] 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.

[43] Robert Serfling.
A Depth Function and a Scale Curve Based on Spatial Quantiles.
In Yadolah Dodge, editor, *Statistical Data Analysis Based on the L1-Norm and Related Methods*, pages 25–38, Basel, 2002. Birkhäuser Basel.

[44] Xiaoping Shi, Yue Zhang, and Yuejiao Fu.
Two-sample tests based on data depth.
Entropy, 25(2), 2023.

[45] Anja Struyf and Peter J Rousseeuw.
Halfspace Depth and Regression Depth Characterize the Empirical Distribution.
Journal of Multivariate Analysis, 69(1):135–153, 1999.

[46] John W. Tukey.
Mathematics and the Picturing of Data.
Proceedings of the International Congress of Mathematicians, Vancouver, 1975, 2:523–531, 1975.

[47] Cédric Villani.
Topics in Optimal Transportation.
Graduate studies in mathematics. American Mathematical Society, 2003.

[48] Yijun Zuo and Xuming He.
On the limiting distributions of multivariate depth-based rank sum statistics and related tests.
The Annals of Statistics, 34(6):2879 – 2896, 2006.

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