Statistical depth: Part III
Miscellanea / Depth in exotic spaces

Stanislav Nagy

Charles University
Department of Probability and Mathematical Statistics

Warsaw 2021

Co-funded by the
Erasmus+ Programme
of the European Union

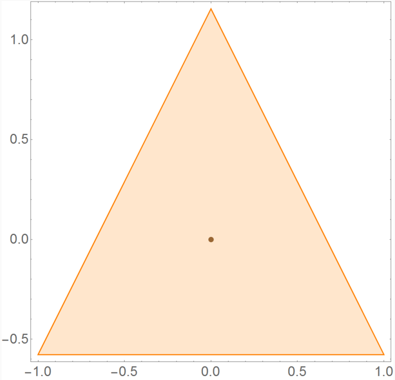
- Miscellanea: Some depth-like procedures
  - Multivariate quantile surfaces
  - Illumination
- Depth in function spaces
- Depth for directional data

Miscellanea: Some depth-like procedures

Definition (Ahidar-Coutrix and Berthet, 2016)
Let $O \in \mathbb{R}^d$, $\delta \in (0, 1/2]$ and $P \in \mathcal{P}\left(\mathbb{R}^d\right)$. The quantile surface $Q(O, \delta)$ of P at level δ about the observer O is the set of all projections of O to the boundaries of all halfspaces $H \in \mathcal{H}$ that satisfy $P(H) = \delta$.

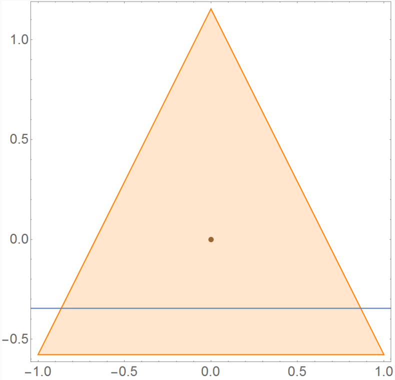
Multivariate quantile surface

Quantile surface of K for $\delta = 0.3$ and O the halfspace median

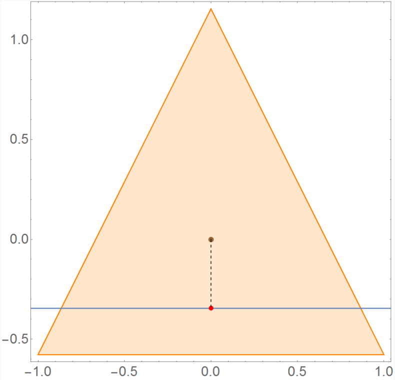


Multivariate quantile surface

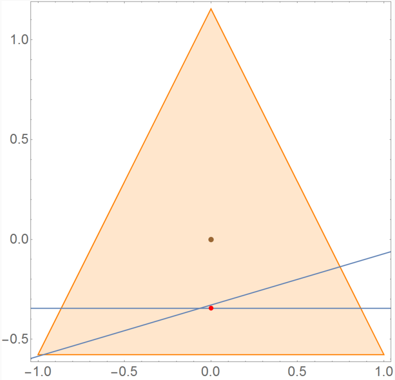
Quantile surface of K for $\delta = 0.3$ and O the halfspace median



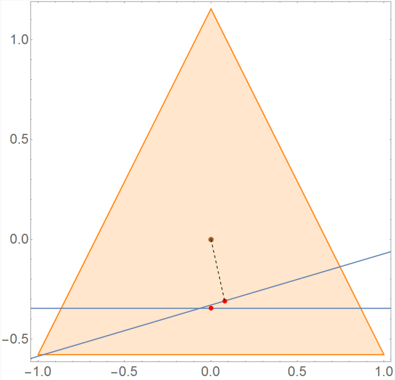
Quantile surface of K for $\delta = 0.3$ and O the halfspace median



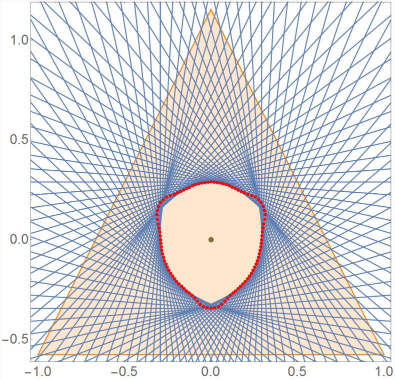
Quantile surface of K for $\delta = 0.3$ and O the halfspace median



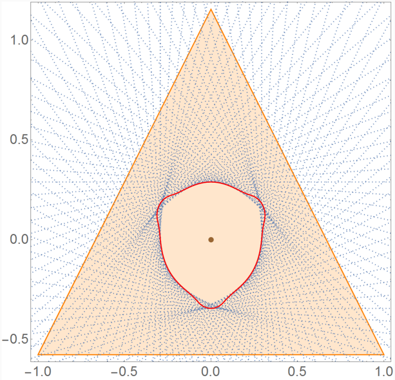
Quantile surface of K for $\delta = 0.3$ and O the halfspace median



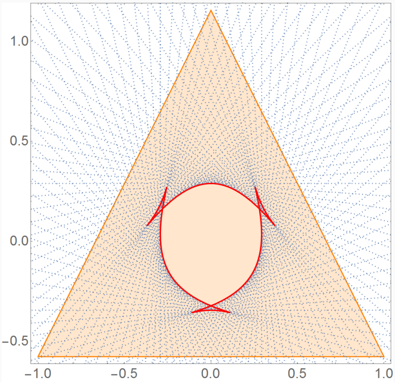
Quantile surface of K for $\delta = 0.3$ and O the halfspace median



Quantile surface of K for $\delta = 0.3$ and O the halfspace median

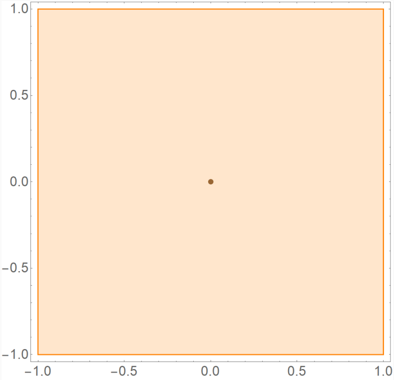


Floating body of K for $\delta = 0.3$

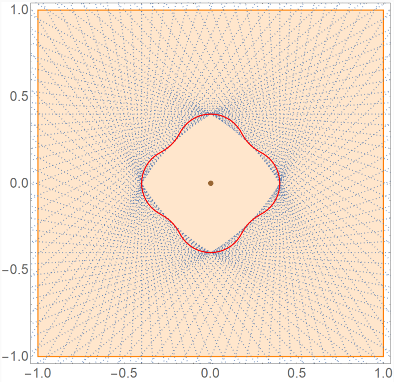


Multivariate quantile surface

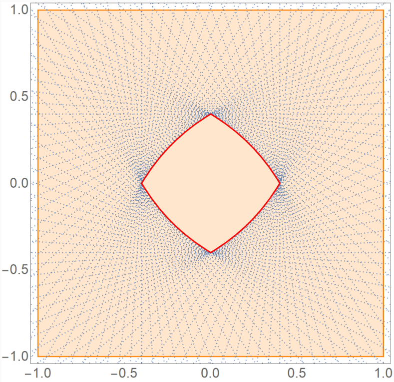
Quantile surface of K for $\delta = 0.3$ and O the halfspace median



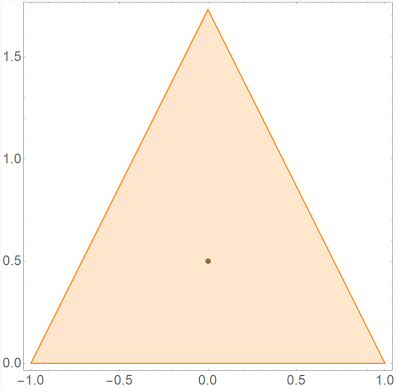
Quantile surface of K for $\delta = 0.3$ and O the halfspace median



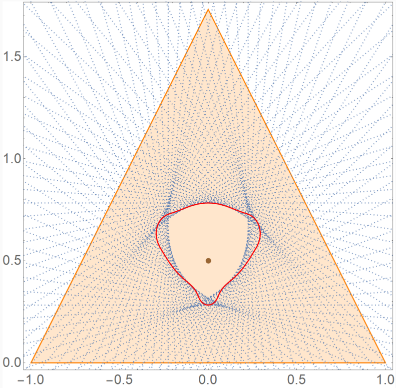
Floating body of K for $\delta = 0.3$



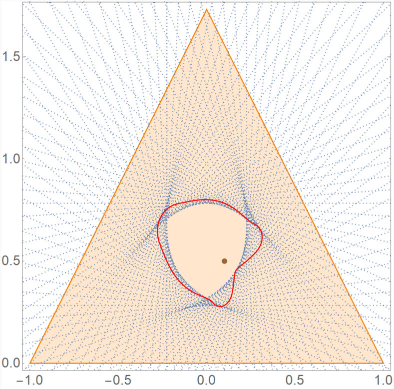
Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



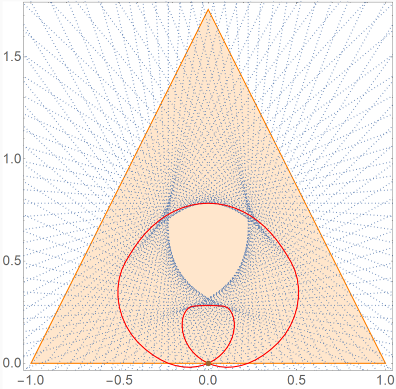
Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



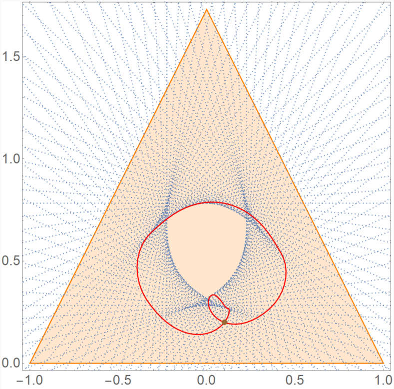
Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



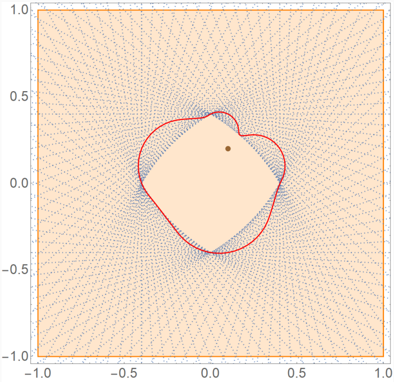
Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



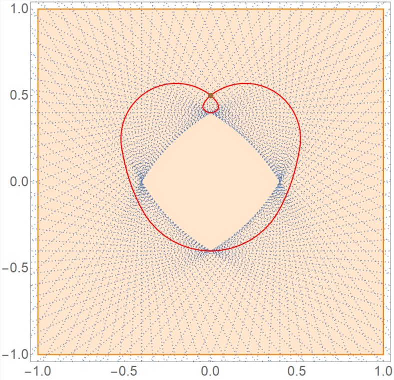
Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



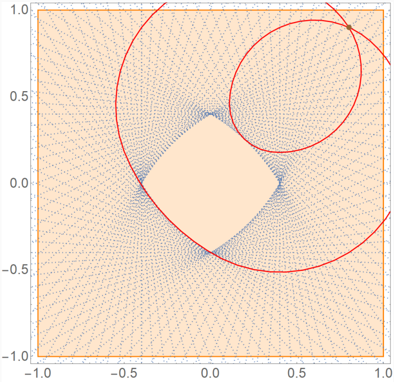
Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



Quantile surface of K for $\delta = 0.3$ and $O \neq$ the halfspace median



Multivariate quantile surface: Properties

Proposition (Ahidar-Coutrix and Berthet, 2016)
Under "minimal" assumptions for O^* the median

- $\{Q(O^*, \delta)\}_{\delta \in (0,1/2]}$ *forms boundaries of an increasing system of embedded star-convex bodies.*
- *uniform strong consistency (in the Hausdorff metric).*
- *uniform weak convergence to a Gaussian process.*
- *uniform law of iterated logarithm.*
- *a Bahadur-Kiefer representation.*
- *non-asymptotic approximation by a Gaussian process.*

Proposition (Ahidar-Coutrix and Berthet, 2016)
For O^* the halfspace median of P, $\{Q(O^*,\delta)\}_{\delta\in(0,1/2]}$ determines the support function of $hD_\delta(P)$.

- Not true in general, holds only if the floating bodies of P exist.
- The relations are more involved — hD_δ is the Wulff shape of the Frank diagram given by $Q(O^*,\delta)$.
  (Ševčovič and Trnovská, 2015)

Proposition
Halfspace depth quantiles $hD_\delta(P)$ are the Wulff shapes corresponding to the integrand

$$\Phi\colon \mathbb{S}^{d-1} \to \mathbb{R}\colon u \mapsto \sup\left\{t\colon \mathrm{P}\left(\langle X, u\rangle \le t\right) \le \delta\right\}.$$

In this representation the multivariate quantile surface $Q(O^*, \delta)$ is given by the anisotropy Φ.

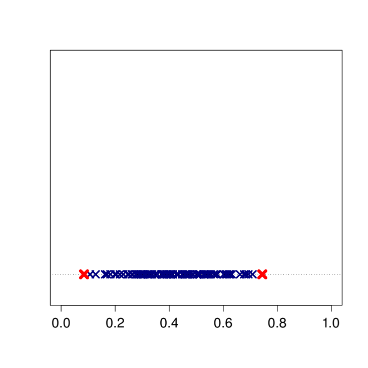
Proposition
Halfspace depth quantiles $hD_\delta(P)$ are the Wulff shapes corresponding to the integrand

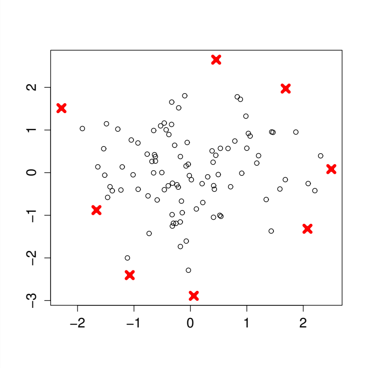
$$\Phi\colon \mathbb{S}^{d-1} \to \mathbb{R}\colon u \mapsto \sup\left\{t\colon \mathrm{P}\left(\langle X, u\rangle \le t\right) \le \delta\right\}.$$

In this representation the multivariate quantile surface $Q(O^*, \delta)$ is given by the anisotropy Φ.

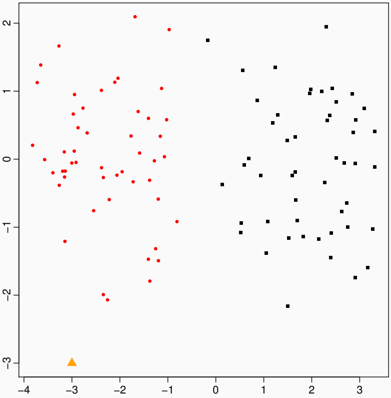
- $hD_\delta(P)$ are "dual" to $Q(O^*, \delta)$ $\overset{?}{\Longrightarrow}$ which properties shown for $Q(O^*, \delta)$ can be transferred? (Ahidar-Coutrix and Berthet, 2016)
- $\overset{?}{\Longrightarrow}$ optimality of the halfspace depth quantiles? (Taylor, 1978)

With increasing dimension d the number of depth-ties increases.

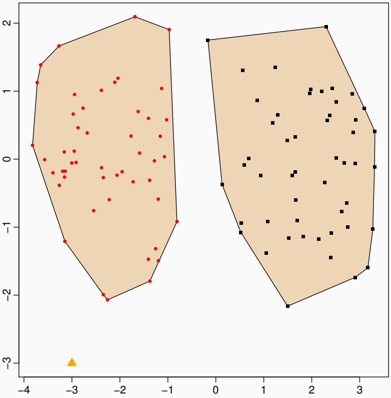




For $X \sim P_1$, $Y \sim P_2$ and $x \sim P_i$, $i \in \{1, 2\}$ unknown, find i.



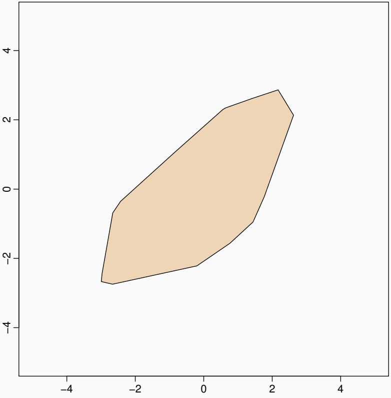
For $X \sim P_1$, $Y \sim P_2$ and $x \sim P_i$, $i \in \{1, 2\}$ unknown, find i.



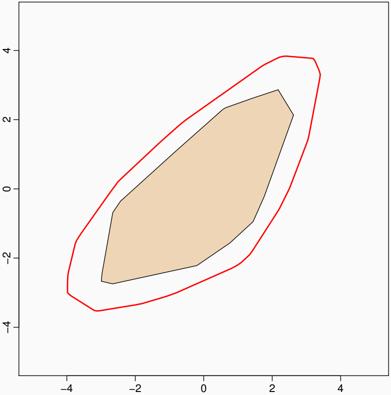
Definition (Werner, 1994)
Let $K \subset \mathbb{R}^d$ be a convex body and $\delta > 0$. The illumination body of K corresponding to δ is given by

$$K^\delta = \left\{ x \in \mathbb{R}^d : \ \mathrm{vol}\left(\mathrm{co}\left(x, K\right)\right) \leq \mathrm{vol}\left(K\right) + \delta \right\}.$$

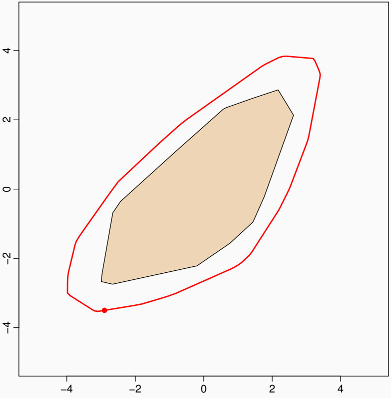
$$K^\delta = \left\{x \in \mathbb{R}^d : \ \mathrm{vol}\left(\mathrm{co}\left(x, K\right)\right) \leq \mathrm{vol}\left(K\right) + \delta\right\}$$



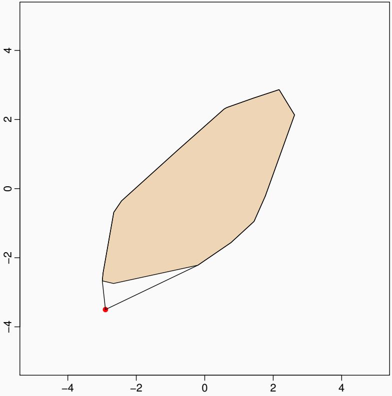
$$K^{\delta} = \left\{x \in \mathbb{R}^d :\ \mathrm{vol}\left(\mathrm{co}\left(x, K\right)\right) \leq \mathrm{vol}\left(K\right) + \delta\right\}$$



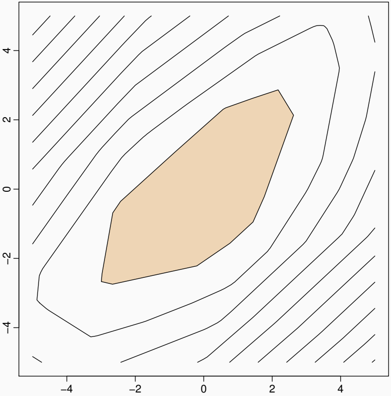
$$K^{\delta} = \left\{x \in \mathbb{R}^d : \ \mathrm{vol}\left(\mathrm{co}\left(x, K\right)\right) \leq \mathrm{vol}\left(K\right) + \delta\right\}$$



$$K^\delta = \left\{x \in \mathbb{R}^d \colon \ \mathrm{vol}\left(\mathrm{co}\left(x, K\right)\right) \leq \mathrm{vol}\left(K\right) + \delta\right\}$$



$$K^{\delta} = \left\{x \in \mathbb{R}^d : \ \mathrm{vol}\left(\mathrm{co}\left(x, K\right)\right) \leq \mathrm{vol}\left(K\right) + \delta\right\}$$



Proposition (Werner 1994, 2006)
It holds that:

- $\left\{K^\delta\right\}_{\delta>0}$ *is an increasing system of concentric convex bodies.*
- *For K an ellipsoid, each K^δ is an ellipsoid of the same shape.*
- *K^δ is invariant w.r.t. rotations.*
- *There exists $b_d > 0$ such that*

$$\Omega(K) = \lim_{\delta\to 0} b_d \frac{\mathrm{vol}\left(K^\delta\right) - \mathrm{vol}\left(K\right)}{\delta^{2/(d+1)}}.$$

Definition
Let $P \in \mathcal{P}\left(\mathbb{R}^d\right)$ and $x \notin \text{co}\left(\mathrm{Supp}\left(P\right)\right)$. The illumination of x w.r.t. P is

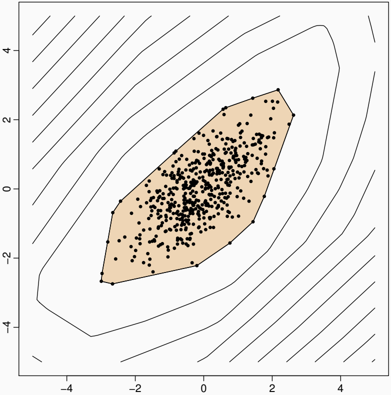
$$\mathscr{I}(x; P) = \text{vol}\left(\text{co}\left(x, \mathrm{Supp}\left(P\right)\right)\right).$$

For $x, y \in \mathbb{R}^d$ such that

$$hD(x; P) = hD(y; P) = 0$$

we say that x is deeper than y if $\mathscr{I}(x; P) < \mathscr{I}(y; P)$.

$$\mathscr{I}(x; P) = \mathrm{vol}\left(\mathrm{co}\left(x, \mathrm{Supp}\left(P\right)\right)\right)$$



Definition (Nagy and Dvořák, 2021)
Let $P \in \mathcal{P}\left(\mathbb{R}^d\right)$, $x \in \mathbb{R}^d$. The robust illumination of x w.r.t. P is

$$\mathscr{I}(x; P) = \operatorname{vol}\left(\operatorname{co}\left(x, \{hD(\cdot; P) \geq (hD(x; P) + s)/2\}\right)\right),$$

where $s = \sup_{y \in \mathbb{R}^d} hD(y; P)$.

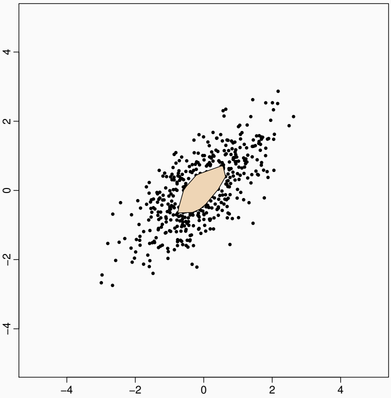
For $x, y \in \mathbb{R}^d$ such that

$$hD(x; P) = hD(y; P)$$

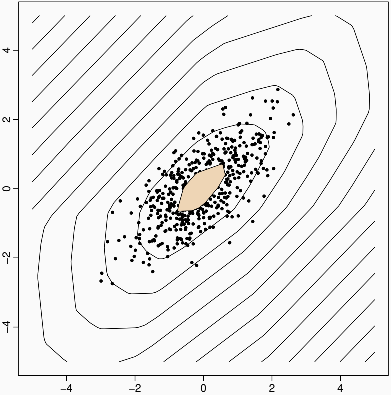
we say that x is deeper than y if $\mathscr{I}(x; P) < \mathscr{I}(y; P)$.

Robust illumination

$$\mathscr{I}(x;P) = \mathrm{vol}\left(\mathrm{co}\left(x, \{hD(\cdot;P) \geq (hD(x;P)+s)/2\}\right)\right)$$

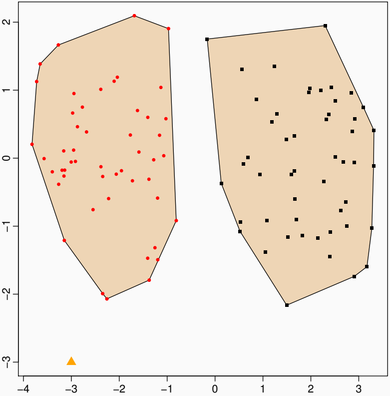


$$\mathscr{I}(x;P) = \mathrm{vol}\left(\mathrm{co}\left(x, \{hD(\cdot;P) \geq (hD(x;P)+s)/2\}\right)\right)$$

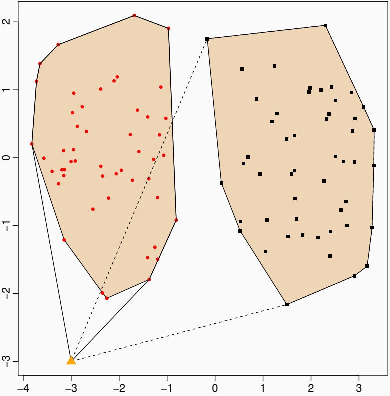


Example: Classification

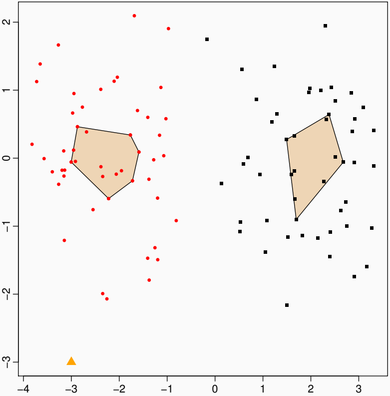
For $X \sim P_1$, $Y \sim P_2$ and $x \sim P_i$, $i \in \{1, 2\}$ unknown, find i.



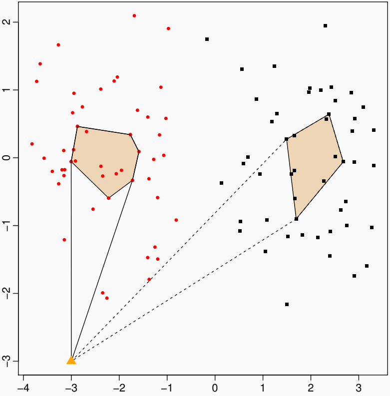
For $X \sim P_1$, $Y \sim P_2$ and $x \sim P_i$, $i \in \{1, 2\}$ unknown, find i.



For $X \sim P_1$, $Y \sim P_2$ and $x \sim P_i$, $i \in \{1, 2\}$ unknown, find i.



For $X \sim P_1$, $Y \sim P_2$ and $x \sim P_i$, $i \in \{1, 2\}$ unknown, find i.



Illumination has an array of good properties:

- duality w.r.t. the halfspace depth,
- conceptual and computational simplicity,
- rotational invariance,
- consistency and robustness,
- invariance for elliptically symmetric distributions,

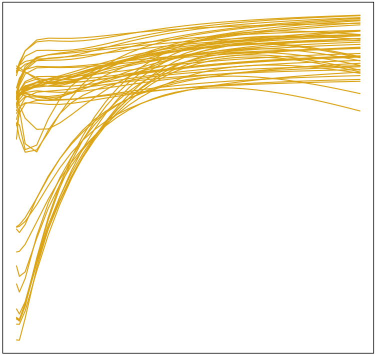
all this with no assumptions on P.

In many applications it outperforms much more complicated methods (Einmahl et al., 2015; Paindaveine and Van Bever, 2013).

Depth in function spaces

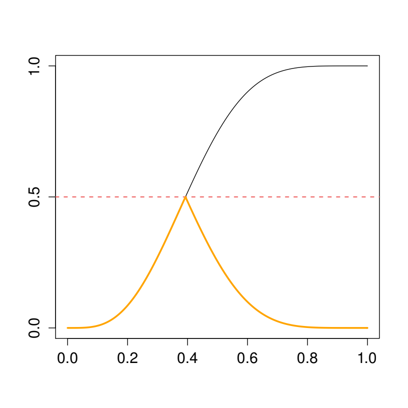
$X \sim P \in \mathcal{P}(\mathcal{F})$ and $X_1, \ldots, X_n$ i.i.d. from P. Consider the depth of functional observations w.r.t. P

$$D \colon \mathcal{F} \times \mathcal{P}(\mathcal{F}) \to [0, 1].$$



Halfspace depth in $\mathbb{R}$

$$hD_1(u;Q) = \min\{\mathrm{F}_Q(u), 1-\mathrm{F}_Q(u-)\} \approx 1/2 - |1/2 - \mathrm{F}_Q(u)|$$



For $\mathcal{F}$ a Banach space and $X \sim P \in \mathcal{P}(\mathcal{F})$, what is the depth?

$$D \colon \mathcal{F} \times \mathcal{P}(\mathcal{F}) \to [0, 1].$$

- For the halfspace depth, only the linear structure of $\mathbb{R}^d$ is needed:

$$hD(x; P) = \inf_{u \in \mathbb{R}^d} P\left(\left\{y \in \mathbb{R}^d \colon \langle y, u\rangle \leq \langle x, u\rangle\right\}\right).$$

- Others, such as the simplicial depth in $\mathbb{R}^d$ depend on d, the dimension of the space.

Depth in function spaces

For $\mathcal{F}$ a Banach space and $X \sim P \in \mathcal{P}(\mathcal{F})$, what is the depth of $x \in \mathcal{F}$?

$$D \colon \mathcal{F} \times \mathcal{P}(\mathcal{F}) \to [0, 1].$$

- Functional halfspace depth: for $\mathcal{F}^*$ the dual space of $\mathcal{F}$

$$hD(x; P) = \inf_{\varphi \in \mathcal{F}^*} P\left(\{y \in \mathcal{F} \colon \varphi(y) \leq \varphi(x)\}\right).$$

- The simplicial depth does not work directly in function spaces.

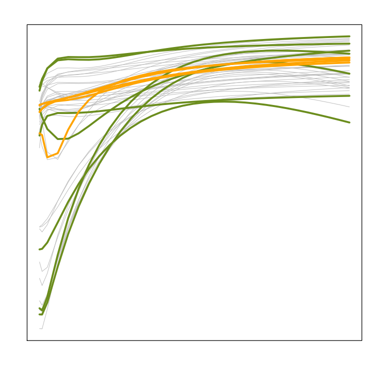
For $\mathcal{F} = L^2(\mathcal{T})$ the space of square-integrable functions, $\mathcal{F}^* = \mathcal{F}$

$$hD(x;P) = \inf_{u \in L^2(\mathcal{T})} P\left(\left\{y \in L^2(\mathcal{T}) : \langle y, u\rangle \leq \langle x, u\rangle\right\}\right)$$

- How to compute the depth?
- What properties does it have?

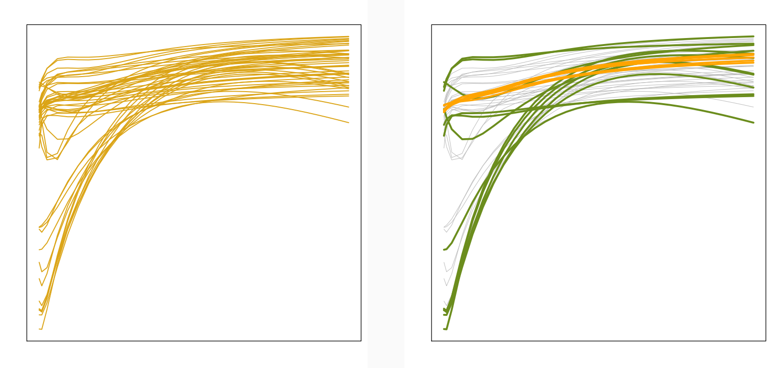
Drawing the directions randomly we obtain the random halfspace depth (Cuesta-Albertos and Nieto-Reyes, 2008)

$$hD_m(x;P) = \min_{u \in \{U_1,\ldots,U_m\}} P\left(\left\{y \in L^2(\mathcal{T}) : \langle y,u \rangle \leq \langle x,u \rangle \right\}\right)$$



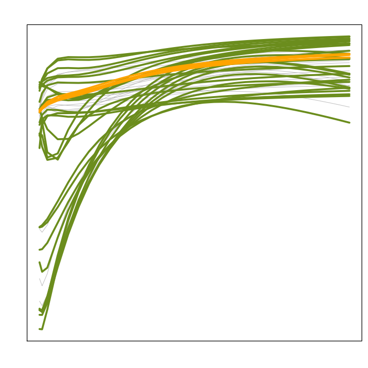
Drawing the directions randomly we obtain the random halfspace depth (Cuesta-Albertos and Nieto-Reyes, 2008)

$$hD_m(x; P) = \min_{u \in \{U_1, \ldots, U_m\}} P\left(\left\{y \in L^2(\mathcal{T}) : \langle y, u \rangle \leq \langle x, u \rangle\right\}\right)$$



Drawing the directions randomly we obtain the random halfspace depth (Cuesta-Albertos and Nieto-Reyes, 2008)

$$hD_m(x;P) = \min_{u \in \{U_1,\ldots,U_m\}} P\left(\left\{y \in L^2(\mathcal{T}) : \langle y,u\rangle \leq \langle x,u\rangle\right\}\right)$$



Random halfspace depth in $L^2(\mathcal{T})$

The random halfspace depth (Cuesta-Albertos and Nieto-Reyes, 2008)

$$hD_m(x; P) = \min_{u \in \{U_1, \dots, U_m\}} P\left(\left\{y \in L^2(\mathcal{T}) : \langle y, u \rangle \leq \langle x, u \rangle\right\}\right)$$

- The depth of a fixed function w.r.t. a fixed measure is random.
- How to choose the number of directions m?
- What distribution to draw from?
- Each functional datum lives in its own dimension!

Each functional datum lives in its own dimension:

Proposition
*For a random sample $X_1, \ldots, X_n$ of **infinite-dimensional** functional data, X_n lies outside of the convex hull of $X_1, \ldots, X_{n-1}$, almost surely.*

- The Hahn-Banach theorem implies that the sample functional halfspace depth is constant zero.
- The (random) halfspace depth necessarily degenerates (as $m \to \infty$).

For (certain) Gaussian processes $P \in \mathcal{P}(\mathcal{F})$ for P-almost all $x \in \mathcal{F}$
(Chakraborty and Chaudhuri, 2013)

$$hD(x; P) = \inf_{\varphi \in \mathcal{F}^*} P(\{y \in \mathcal{F} \colon \varphi(y) \leq \varphi(x)\}) = 0.$$

Many other functional depths (López-Pintado and Romo, 2009, 2011; Zuo and Serfling, 2000) degenerate too.

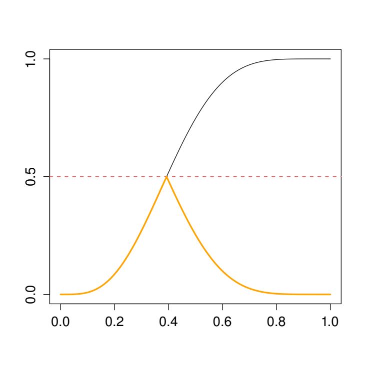
Condition 0: Depth should not degenerate. That is, it is not allowed that $D(x; P) = 0$ for P-almost all $x \in \mathcal{F}$.

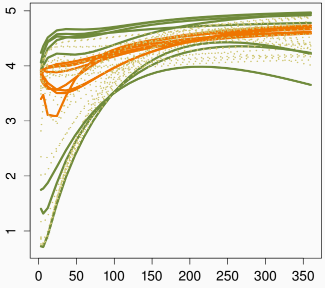
➔ Restrict the set of projections in hD from the dual $\mathcal{F}^*$ to a smaller, but still representative and well interpretable subset.

Average depth of a functional value

(Fraiman and Muniz, 2001; Cuevas and Fraiman, 2009)

$$FD\left(x; P\right) = \int_{\mathcal{T}} D_1(x(t), P_t) \, \mathrm{d}\, t, \qquad D_1(u; Q) = 1/2 - \left|1/2 - \mathrm{F}_Q(u)\right| .$$

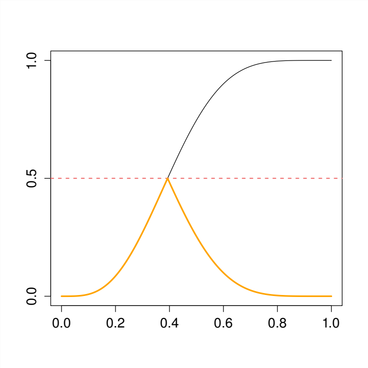


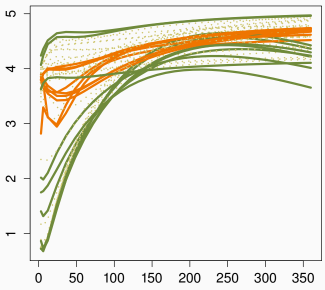


Smallest depth of a functional value

(Mosler, 2013; Narisetty and Nair, 2016)

$$ID\,(x;P) = \inf_{t\in\mathcal{T}} D_1\,(x(t);P_t)\,, \qquad D_1(u;Q) = 1/2 - |1/2 - \mathrm{F}_Q(u)|\,.$$





Basic types of depth for functional data:

- integrated depth

$$FD\,(x;P) = \int_{\mathcal{T}} D_1(x(t), P_t)\,\mathrm{d}\,t,$$

- infimal depth

$$ID\,(x;P) = \inf_{t\in\mathcal{T}} D_1\,(x(t);P_t)\,.$$

For a Banach space B and B^* its dual, $P \in \mathcal{P}(B)$, $\Phi \subset B^*$:

- integrated depth

$$FD(x;P) = \int_{\Phi} D_1(\varphi(x), P_{\varphi(X)})\, \mathrm{d}\,\lambda(\varphi),$$

- infimal depth

$$ID(x;P) = \inf_{\varphi \in \Phi} D_1\left(\varphi(x), P_{\varphi(X)}\right).$$

The set $\Phi \subset B^*$ is typically the collection of evaluation functionals

$$\{\varphi_t \colon x \mapsto x(t) \colon t \in \mathcal{T}\},$$

but not necessarily so. λ is a measure on Φ.

Proposition
The integrated depth does not degenerate, but the infimal depth "almost" does.

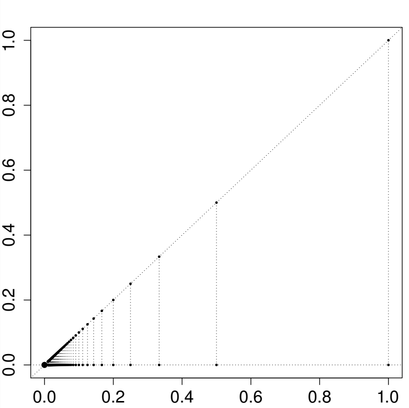
Example: Consider $X \sim P \in \mathcal{P}(\mathcal{C})$ given as a linear interpolant of

- $X(0) = 0$, and
- $X(1/n) = \mathrm{Bernoulli}(1/2)/n$ independent for $n = 1, 2, \ldots$.

Then $ID(x; P_n) = 0$ for all $x \in \mathcal{C}$, almost surely.

For $X \sim P$ the randomly jumping function

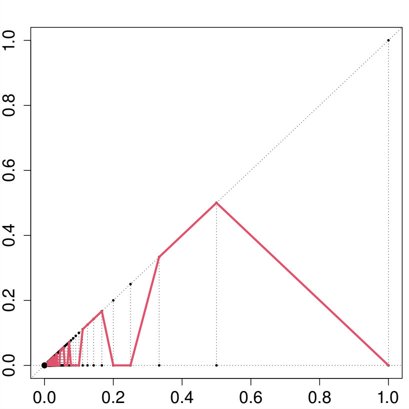
$$ID(x; P) = 1/2 \times \mathbb{I}\left[0 \leq x(t) \leq t \text{ for all } t \in [0,1]\right]$$



Infimal depths: Degeneracy problem

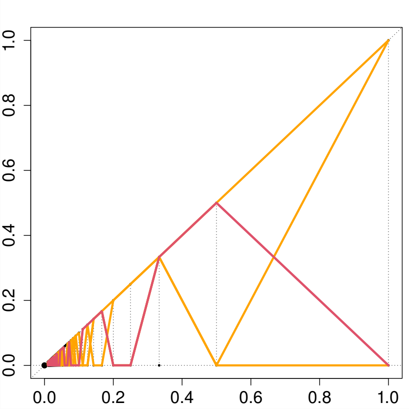
For $X_1, \ldots, X_n$ a random sample from P with empirical measure P_n

$$ID(x; P_n) = 0 \text{ for any } x \neq X_i,\ i = 1, \ldots, n.$$



For $X_1, \ldots, X_n$ a random sample from P with empirical measure P_n

$$ID(x; P_n) = 0 \text{ for any } x \neq X_i, \, i = 1, \ldots, n.$$



Depth distribution

Consider the depth distribution of $x \in L^2(\mathcal{T})$, that is the law of

$$D_x \colon \left(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda\right) \to [0,1] \colon t \mapsto hD\left(x(t); P_t\right)$$

being a random variable on $\mathcal{T}$.

- The integrated depth is the mean of D_x

$$FD(x; P) = \int_{\mathcal{T}} hD\left(x(t); P_t\right) \, \mathrm{d}\,\lambda(t) = \mathsf{E}\, D_x.$$

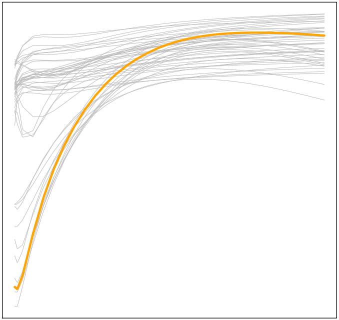
- The infimal depth is the (essential) infimum of D_x

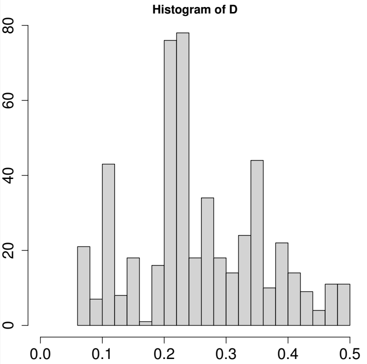
$$ID(x; P) = \inf_{t \in \mathcal{T}} hD\left(x(t); P_t\right),$$

  that is the lower end-point of the support of D_x.

The depth distribution of $x \in L^2(\mathcal{T})$ w.r.t. the random sample

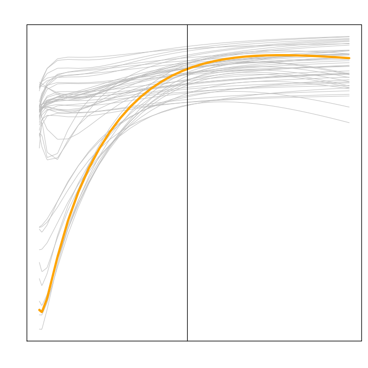
$$D_x\colon\ (\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda) \to [0,1]\colon\ t \mapsto hD(x(t); P_t)$$

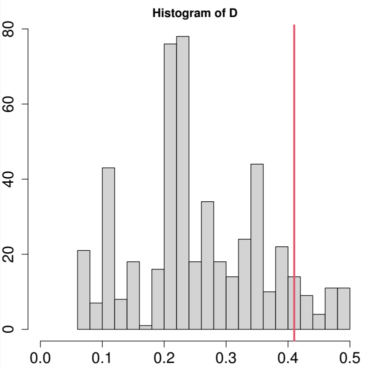




The depth distribution of $x \in L^2(\mathcal{T})$ w.r.t. the random sample

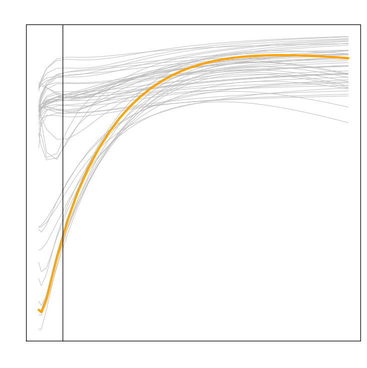
$$D_x\colon\ (\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda) \to [0,1]\colon\ t \mapsto hD(x(t); P_t)$$

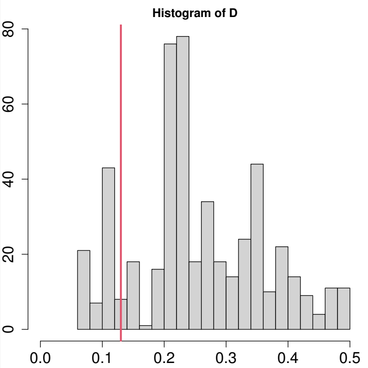




The depth distribution of $x \in L^2(\mathcal{T})$ w.r.t. the random sample

$$D_x \colon \left(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda\right) \to [0,1] \colon t \mapsto hD\left(x(t); P_t\right)$$





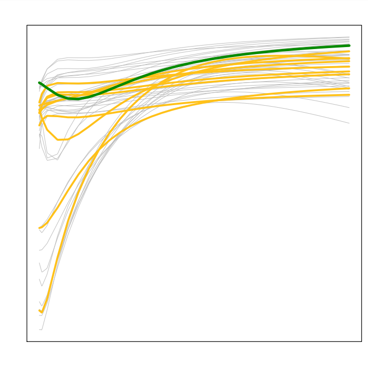
The K-integrated depth with $K \in \mathbb{R}$

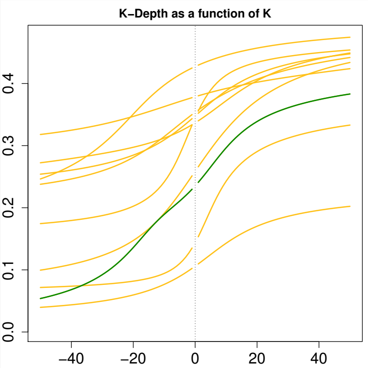
$$
\begin{aligned}
D^K(x;P) &= \left(\int_{\mathcal{T}} (hD(x(t);P_t) + 1/2)^k \,\mathrm{d}\,\lambda(t)\right)^{1/k} - 1/2 \\
&= \left(\mathsf{E}\,(D_x + 1/2)^k\right)^{1/k} - 1/2
\end{aligned}
$$

is, basically, the k-th moment of the depth distribution of x. We obtain a family of depths

- for $k = 1$ the usual integrated depth;
- as $k \to -\infty$ a version of the infimal depth;
- choice of k allows us to fine tune.

The trajectories $K \mapsto D^K(x;P) = \left(\mathrm{E}\left(D_x + 1/2\right)^k\right)^{1/k} - 1/2$





General functional depths?

One can choose any (location) parameter L of the depth distribution

$$D_L(x; P) = L(D_x)$$

to obtain a custom tailored depth functional. Examples are

- quantiles,
- trimmed means,
- M-estimators...

The resulting depths possess quite different properties.

Case in point: **Sample version consistency**.

A theoretical issue: Consistency

Let $P_n \in \mathcal{P}(B)$ be the (random) empirical measure of a random sample $X_1, \ldots, X_n$ from P.

A depth D on space B is

- consistent if $D(x; P_n) \xrightarrow[n\to\infty]{\text{a.s.}} D(x; P)$ for all $x \in B$;
- uniformly consistent if $\sup_{x\in B} |D(x; P_n) - D(x; P)| \xrightarrow[n\to\infty]{\text{a.s.}} 0$.

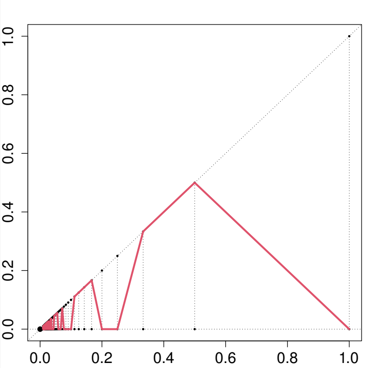
➔ In $B = \mathbb{R}^d$, the halfspace depth is uniformly consistent.

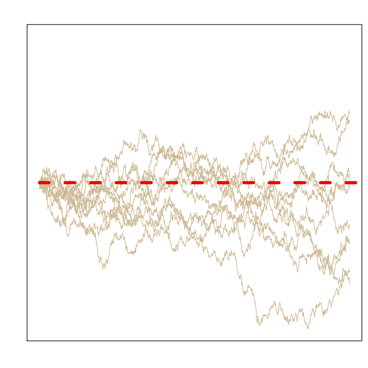
➔ In function spaces uniform consistency requires new theories.

➔ Functional depths are often not consistent uniformly.

ID is not consistent for, e.g., *P* the Wiener measure

(Gijbels and Nagy, 2015)





Integrated-type depths are uniformly consistent

By the dominated convergence theorem (and measurability)

$$
\begin{aligned}
&\sup_{x \in L^2(\mathcal{T})} \left| FD\,(x; P) - FD\,(x; P_n) \right| \\
&= \sup_{x \in L^2(\mathcal{T})} \left| \int_0^1 hD\,(x(t); P_t) - hD\,(x(t); P_{n,t})\, \mathrm{d}\, t \right| \\
&\leq \int_0^1 \sup_{x \in L^2(\mathcal{T})} \left| hD\,(x(t); P_t) - hD\,(x(t); P_{n,t}) \right| \mathrm{d}\, t \\
&\leq \int_0^1 \sup_{u \in \mathbb{R}} \left| hD\,(u; P_t) - hD\,(u; P_{n,t}) \right| \mathrm{d}\, t, \\
&\quad \xrightarrow[n \to \infty]{\text{a.s.}} 0
\end{aligned}
$$

FD is universally consistent, but this proof is not complete.
(Nagy et al., 2016)

Problem: Uniform consistency

The Vapnik-Červonenkis theory gives that

$$\sup_{u\in\mathbb{R}} \left| hD\left(u; P_t\right) - hD\left(u; P_{n,t}\right)\right| \xrightarrow[n\to\infty]{\text{a.s.}} 0$$

for each $t \in \mathcal{T}$ separately. That is, the convergence is true for all $\omega \notin N_t$ with $\mathrm{P}(N_t) = 0$. But,

- There are uncountably many different $t \in \mathcal{T}$ and uncountably many such sets $N_t \subset \Omega$.
- The union $\bigcup_{t\in\mathcal{T}} N_t$ does not have to be a P-null set.

To conclude consistency, that is

$$\int_0^1 \sup_{u\in\mathbb{R}} \left| hD\left(u; P_t\right) - hD\left(u; P_{n,t}\right)\right| \mathrm{d}\, t \xrightarrow[n\to\infty]{\text{a.s.}} 0,$$

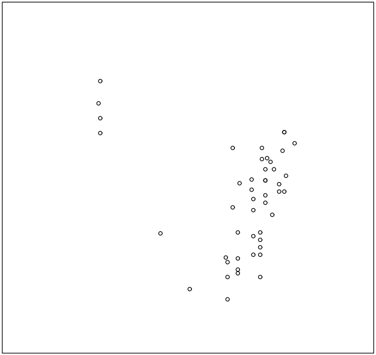
one needs to guarantee that $\mathrm{P}\left(\bigcup_{t\in\mathcal{T}} N_t\right) = 0$.

Theorem
All integrated depths are uniformly consistent over $L^2(\mathcal{T})$, for any $P \in \mathcal{P}\left(L^2(\mathcal{T})\right)$.

- Proof uses measurability / abstract Fubini's theorem (Nagy et al., 2016; 2021).
- For general functional depths with location parameters L, it is not easy to establish $\mathsf{P}\left(\bigcup_{t\in\mathcal{T}} N_t\right) = 0$.

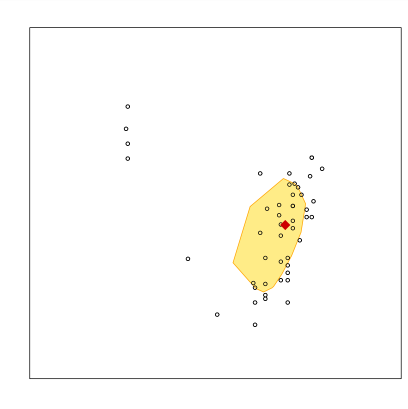
Bagplot — depth-based boxplot for multivariate data

(Rousseeuw et al., 1999)



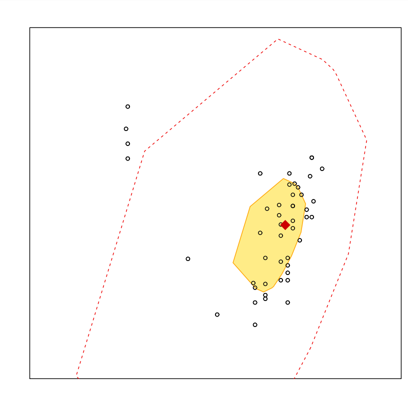
Bagplot — depth-based boxplot for multivariate data

(Rousseeuw et al., 1999)



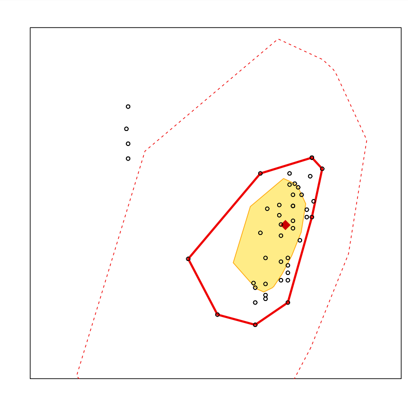
Bagplot — depth-based boxplot for multivariate data

(Rousseeuw et al., 1999)



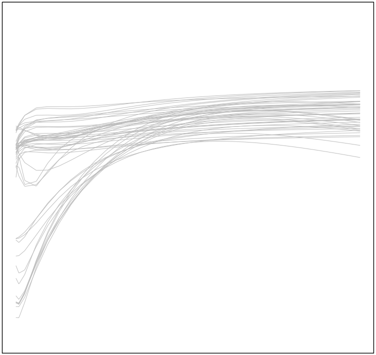
Bagplot — depth-based boxplot for multivariate data

(Rousseeuw et al., 1999)

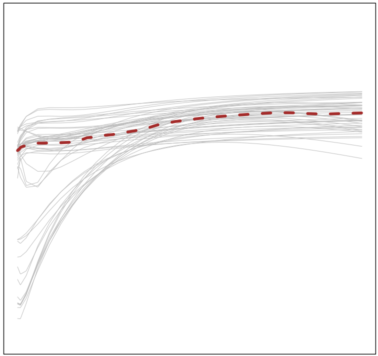


Functional boxplot

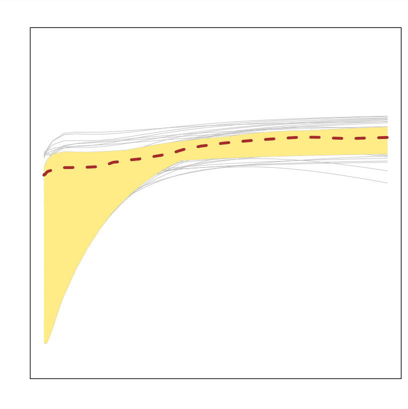
Functional boxplot based on integrated depths for functional data (Sun and Genton, 2011)



Functional boxplot based on integrated depths for functional data (Sun and Genton, 2011)

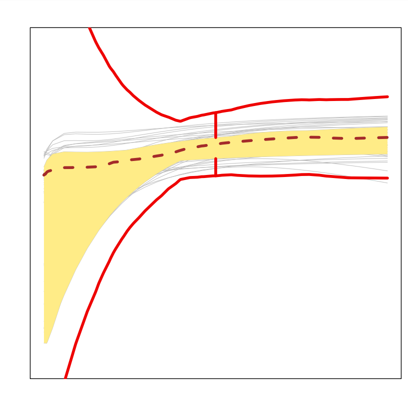


Functional boxplot based on integrated depths for functional data (Sun and Genton, 2011)

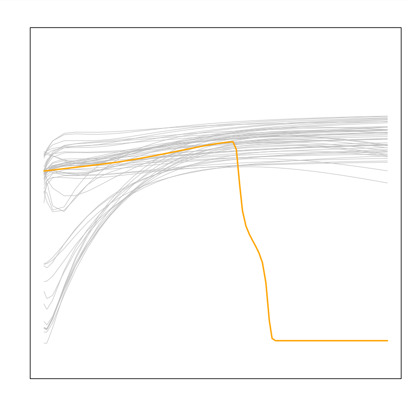


Functional boxplot

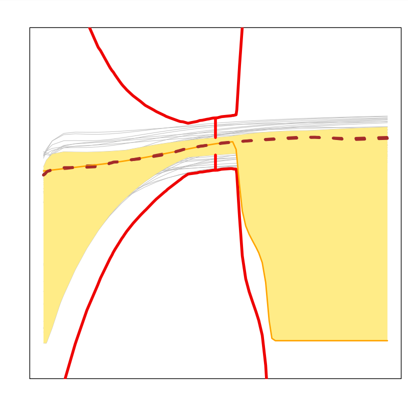
Functional boxplot based on integrated depths for functional data (Sun and Genton, 2011)



Functional boxplot based on integrated depths for functional data (Sun and Genton, 2011)



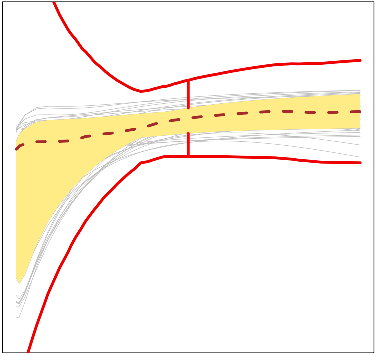
Functional boxplot based on integrated depths for functional data (Sun and Genton, 2011)



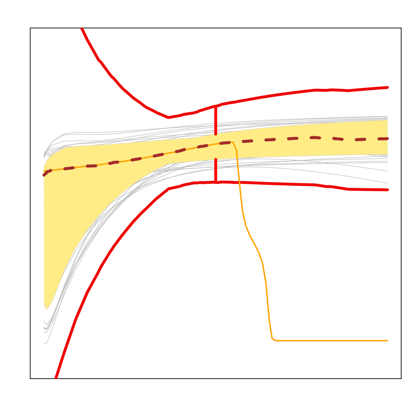
Functional boxplot based on integrated depths for functional data (Sun and Genton, 2011):

- When based on integrated depths, they fail to be robust.
- No guarantees for the nominal coverage probability of the box.
- What is the population version of the functional boxplot?

Boxplot based on an infimal depth (Narisetty and Nair, 2016)



Boxplot based on an infimal depth (Narisetty and Nair, 2016)

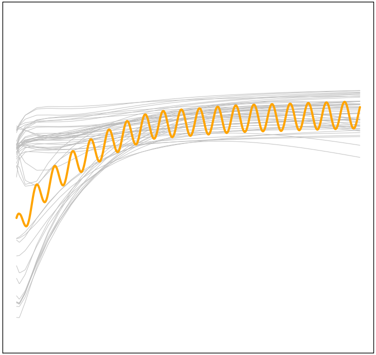


Functional boxplot

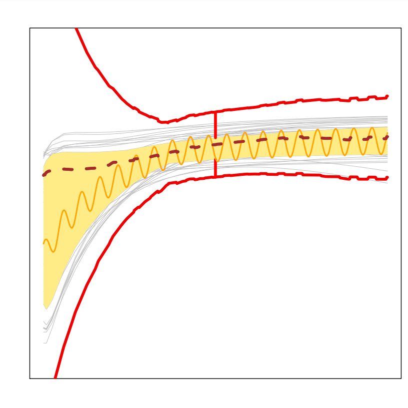
Boxplots and central regions based on an infimal depth (Narisetty and Nair, 2016) are useful also in other contexts:

- Envelope testing (Ripley, 1977; Myllymäki et al., 2016);
- Functional ANOVA (Mrkvička et al., 2020);
- Confidence/prediction regions for functional data (Diquigiovanni et al, 2021);
- Analysis of functional records (Martínez-Hernández and Genton, 2019);
- ...

Blind to the shapes of the functions and phase variation



Boxplots are blind to the shapes of the functions



Blindness to shape features

Boxplots/infimal depths are blind to the shapes of the functions and phase variation:

- The reason being the univariate nature of the depth

$$ID(x; P) = \inf_{t \in \mathcal{T}} D_1(x(t); P_t).$$

- Impossible to be avoided using bands directly.
- Are bands really analogues of convex hulls for functional data?
- How to better visualize functional data?

Conclusions

What we know:

- Functional depth is a very active field of FDA,
- with many potential applications,
- and many depths have been proposed.
- The selection of a depth is crucial.

Open problems:

- Desiderata for the depth?
- Statistical properties.
- How to choose a depth?
- Efficient visualization of functional data? Are bands the way to go?
- Which depths characterize distributions? (Wynne and Nagy, 2021+)

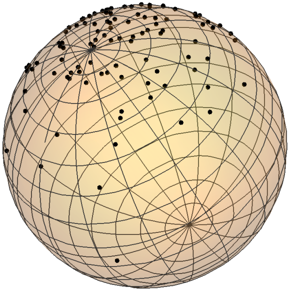
Depth for directional data

Directional data: Angular halfspace depth

Directional data means $P \in \mathcal{P}\left(\mathbb{S}^{d-1}\right)$ (Ley and Verdebout, 2017).

The angular halfspace depth (Small, 1987; Liu and Singh, 1992) of a point $x \in \mathbb{S}^{d-1}$ w.r.t. P

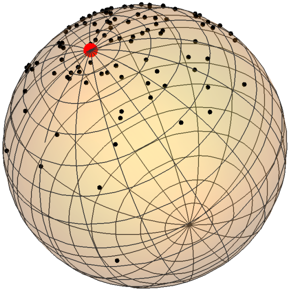
$$ahD\left(x;P\right) = \inf\left\{P\left(H\right) : H \in \mathcal{H}_0 \text{ and } x \in H\right\}.$$



Directional data means $P \in \mathcal{P}\left(\mathbb{S}^{d-1}\right)$ (Ley and Verdebout, 2017).

The angular halfspace depth (Small, 1987; Liu and Singh, 1992) of a point $x \in \mathbb{S}^{d-1}$ w.r.t. P

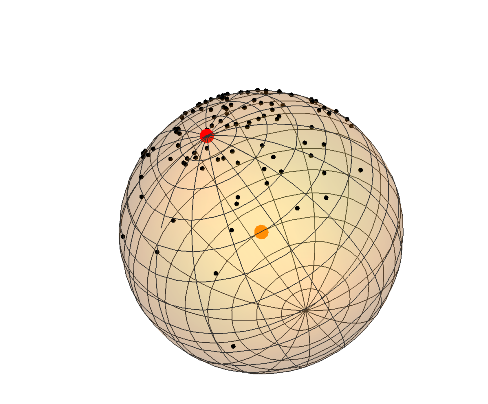
$$ahD\,(x; P) = \inf \left\{P\,(H) : H \in \mathcal{H}_0 \text{ and } x \in H\right\}.$$



Directional data means $P \in \mathcal{P}\left(\mathbb{S}^{d-1}\right)$ (Ley and Verdebout, 2017).

The angular halfspace depth (Small, 1987; Liu and Singh, 1992) of a point $x \in \mathbb{S}^{d-1}$ w.r.t. P

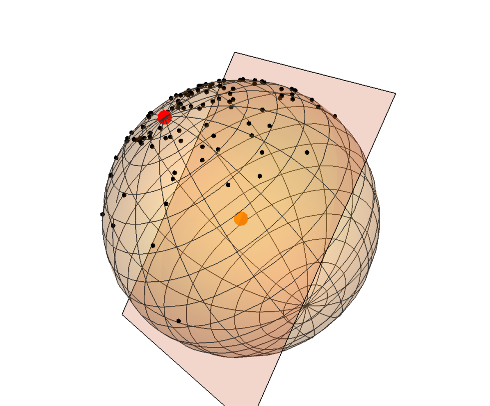
$$ahD\,(x; P) = \inf\left\{P\,(H) : H \in \mathcal{H}_0 \text{ and } x \in H\right\}.$$



Directional data means $P \in \mathcal{P}\left(\mathbb{S}^{d-1}\right)$ (Ley and Verdebout, 2017).

The angular halfspace depth (Small, 1987; Liu and Singh, 1992) of a point $x \in \mathbb{S}^{d-1}$ w.r.t. P

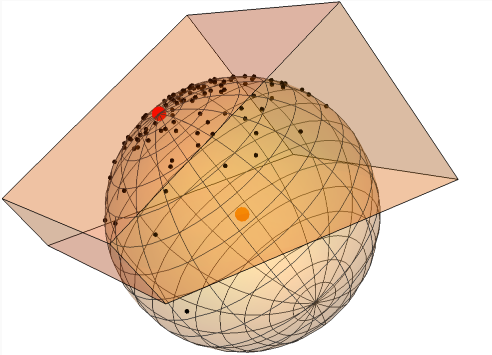
$$ahD(x; P) = \inf\left\{P(H) : H \in \mathcal{H}_0 \text{ and } x \in H\right\}.$$



Directional data means $P \in \mathcal{P}\left(\mathbb{S}^{d-1}\right)$ (Ley and Verdebout, 2017).

The angular halfspace depth (Small, 1987; Liu and Singh, 1992) of a point $x \in \mathbb{S}^{d-1}$ w.r.t. P

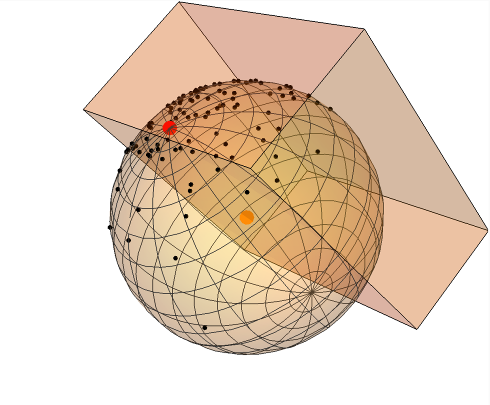
$$ahD\,(x; P) = \inf \left\{P\,(H) : H \in \mathcal{H}_0 \text{ and } x \in H\right\}.$$



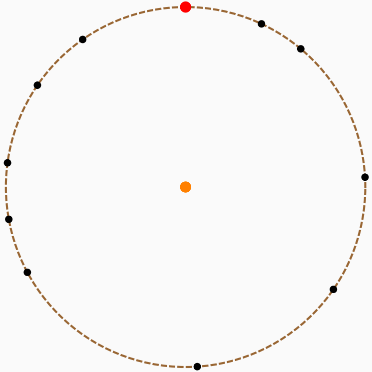
Directional data means $P \in \mathcal{P}\left(\mathbb{S}^{d-1}\right)$ (Ley and Verdebout, 2017).

The angular halfspace depth (Small, 1987; Liu and Singh, 1992) of a point $x \in \mathbb{S}^{d-1}$ w.r.t. P

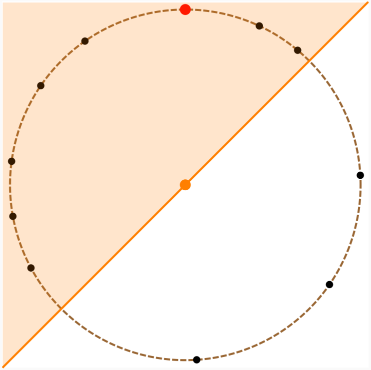
$$ahD\,(x; P) = \inf\left\{P\,(H) : H \in \mathcal{H}_0 \text{ and } x \in H\right\}.$$



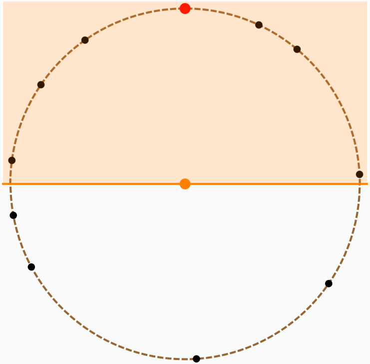
$$ahD\left(x; P_n\right) = \min \frac{\text{\# of observations in } H \in \mathcal{H}_0 \text{ that contains } x}{n}$$



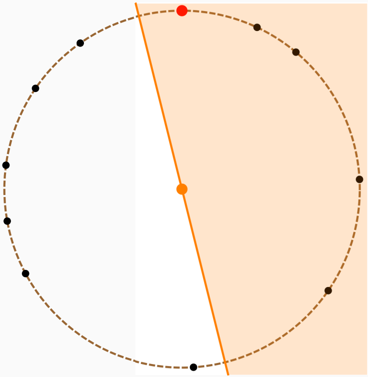
$$ahD\,(x; P_n) = \min \frac{\text{\# of observations in } H \in \mathcal{H}_0 \text{ that contains } x}{n}$$



$$ahD\,(x; P_n) = \min \frac{\#\text{ of observations in } H \in \mathcal{H}_0 \text{ that contains } x}{n}$$



$$ahD\left(x; P_n\right) = \min \frac{\text{\# of observations in } H \in \mathcal{H}_0 \text{ that contains } x}{n}$$



Known theory for *ahD* is quite analogous to that of *hD*:

- rotation invariance and sample version consistency, or
- quasi-concavity similarly as for *hD*.
- Distinctive is the existence of a hemisphere of minimum depth — an open hemisphere $S \subset \mathbb{S}^{d-1}$ with

$$ahD\,(x; P) = \inf\left\{P(H) \colon H \in \mathcal{H}_0\right\} \text{ for all } x \in S.$$

But, the theory is less developed. Not much is known about e.g.

- asymptotic normality of the sample version,
- asymptotics and the convergence of the level sets,
- statistical robustness, or
- **algorithms**.

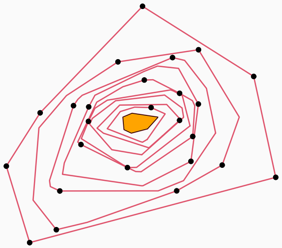
(Linear) Halfspace depth: Computation

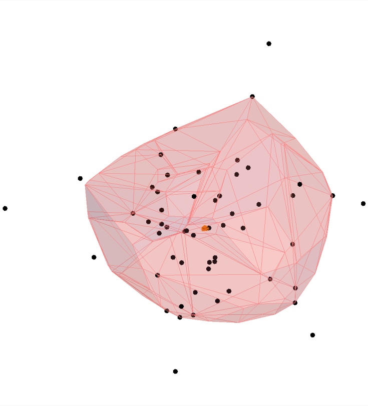
Computational aspects of hD:

- determining $hD(x; P)$ exactly is in general NP-hard (Johnson and Preparata, 1978);
- reasonably fast exact algorithms are available for low $d \leq 5$ (Rousseeuw and Struyf, 1998; Dyckerhoff and Mozharovskyi, 2016);
- very fast approximate algorithms exist (Dyckerhoff, 2004; Chen et al., 2013; Dyckerhoff et al., 2021);
- fast computation of central regions / halfspace median (Liu et al., 2019).

Implemented in R packages `depth` (Genest et al., 2008), `ddalpha` (Pokotylo et al., 2013), `TukeyRegion` (Barber and Mozharovskyi, 2017), or `mrfDepth` (Segaert et al., 2017).

Datasets in $\mathbb{R}^2$ (left) and $\mathbb{R}^3$ (right) with central regions and medians





Quoting Pandolfo, Paindaveine, and Porzio (2018, p. 594)

> *"The main drawback of […] the angular halfspace depth is the computational effort it requires, especially for higher dimensions d."*

The only implementation `sdepth` in package `depth` (Genest et al., 2008) in R allows just $d = 2, 3$.

Note: Here, $P \in \mathcal{P}\left(\mathbb{S}^{d-1}\right)$ is always absolutely continuous.

For $e_d = (0, \ldots, 0, 1)$ we denote

$$\mathbb{S}_+^{d-1} = \left\{x \in \mathbb{S}^{d-1} \colon \langle x, e_d \rangle > 0\right\}, \quad \mathbb{S}_-^{d-1} = \left\{x \in \mathbb{S}^{d-1} \colon \langle x, e_d \rangle < 0\right\},$$

the "northern" and the "southern" hemisphere, and write

$$G = \left\{x \in \mathbb{R}^d \colon \langle x, e_d \rangle = 1\right\}$$

for the "horizontal" hyperplane that touches $\mathbb{S}^{d-1}$ at e_d.

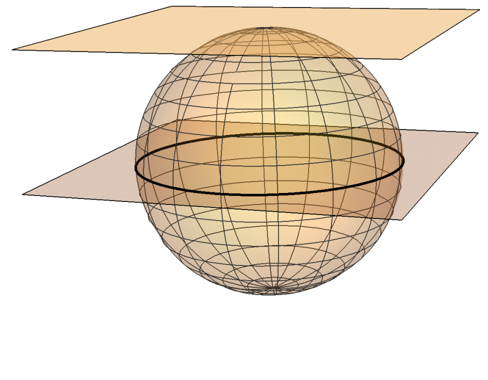
The gnomonic projection of $\mathbb{S}^{d-1}$ maps $x \in \mathbb{S}_+^{d-1}$ to

$$\pi(x) = x / \langle x, e_d \rangle \in G.$$

For $x \in \mathbb{S}_-^{d-1}$ we define $\pi(x) = \pi(-x)$.

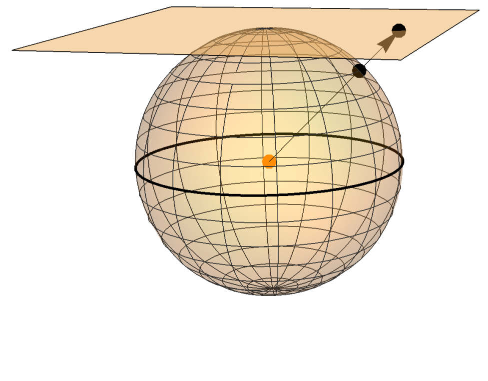
The gnomonic projection

$$x \in \mathbb{S}_+^{d-1} \mapsto \pi(x) = x/\langle x, e_d \rangle \in G.$$



The gnomonic projection

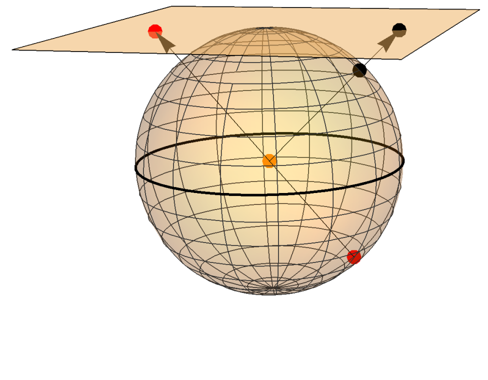
$$x \in \mathbb{S}^{d-1}_{+} \mapsto \pi(x) = x / \langle x, e_d \rangle \in G.$$



Idea: Gnomonic projection and halfspaces

The gnomonic projection

$$x \in \mathbb{S}_+^{d-1} \mapsto \pi(x) = x / \langle x, e_d \rangle \in G.$$

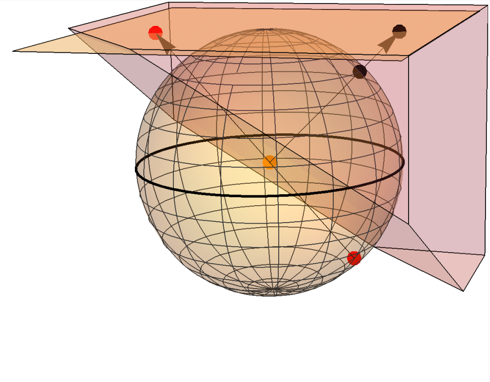


Toward signed halfspace depth I

For any $H \in \mathcal{H}_0$ it holds true that

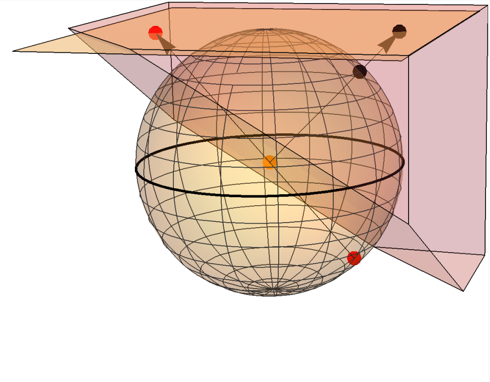
$$\pi\left(H \cap \mathbb{S}_+^{d-1}\right) = H \cap G, \quad \pi\left(H \cap \mathbb{S}_-^{d-1}\right) = G \setminus \operatorname{int}(H),$$

where $\operatorname{int}(H)$ is the interior of H.



We define a signed measure $P_{\pm}$ on G by

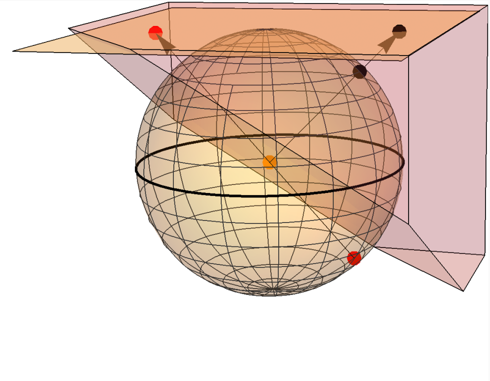
$$P_{\pm}(H \cap G) = P\left(H \cap \mathbb{S}_{+}^{d-1}\right) - P\left(\mathbb{S}_{-}^{d-1} \setminus H\right) \qquad \text{for } H \in \mathcal{H}_0.$$



Altogether for any $H \in \mathcal{H}_0$ and $x \in \mathbb{S}_+^{d-1}$

$$P(H) = P\left(\mathbb{S}_-^{d-1}\right) + P_\pm (H \cap G),$$

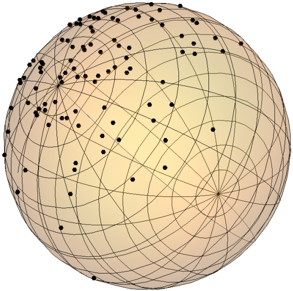
$$ahD(x;P) = P\left(\mathbb{S}_-^{d-1}\right) + \inf \{P_\pm (H) : H \in \mathcal{H} \text{ and } x \in H\}.$$



Computation of $ahD(x;P)$ in $\mathbb{S}^{d-1}$ is equivalent with the evaluation of the signed halfspace depth in $\mathbb{R}^{d-1}$

$$\inf\left\{P_{\pm}(H) : H \in \mathcal{H} \text{ and } x \in H\right\}.$$

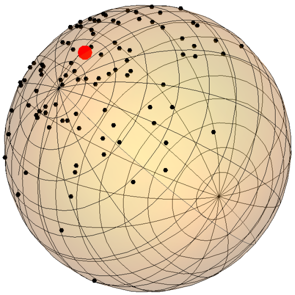
For the latter, fast algorithms for hD can be adapted.



Computation of $ahD(x;P)$ in $\mathbb{S}^{d-1}$ is equivalent with the evaluation of the signed halfspace depth in $\mathbb{R}^{d-1}$

$$\inf\left\{P_{\pm}(H) : H \in \mathcal{H} \text{ and } x \in H\right\}.$$

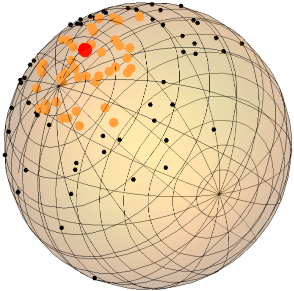
For the latter, fast algorithms for *hD* can be adapted.



Computation of $ahD(x; P)$ in $\mathbb{S}^{d-1}$ is equivalent with the evaluation of the signed halfspace depth in $\mathbb{R}^{d-1}$

$$\inf \{P_{\pm}(H) : H \in \mathcal{H} \text{ and } x \in H\}.$$

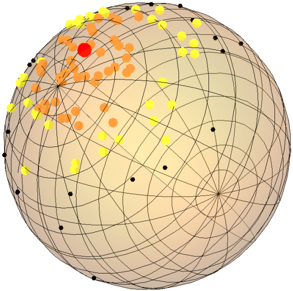
For the latter, fast algorithms for *hD* can be adapted.



Computation of $ahD(x; P)$ in $\mathbb{S}^{d-1}$ is equivalent with the evaluation of the signed halfspace depth in $\mathbb{R}^{d-1}$

$$\inf \{P_{\pm}(H) : H \in \mathcal{H} \text{ and } x \in H\}.$$

For the latter, fast algorithms for *hD* can be adapted.



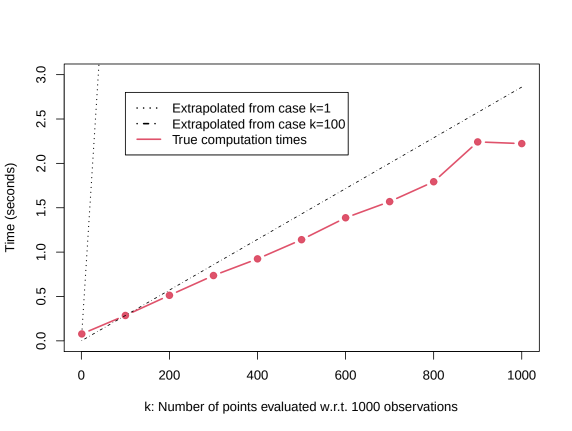
Comparison with `sdepth` from the R package `depth`

n	10	50	100	500	1000	5000
fast	0.000018	0.00015	0.0006	0.020	0.080	2.34
`sdepth`	0.001440	0.11200	0.8600	109.310	944.620	—
ratio	80	747	1433	5466	11808	—

Table 1: Computation times (in seconds) of $ahD\,(x; P)$ for a single point x w.r.t. a random sample of n observations in dimension $d = 3$. In the bottom row the fraction `sdepth`/fast.

Computation times II: Scalability

Computing the depth of k points w.r.t. a dataset of size 1000.



Compare with $\sim$ 944 seconds for *ahD* of a single observation w.r.t. a dataset of size 1000 using `sdepth`.

Computation of *ahD* in $\mathbb{S}^{d-1}$ is not that hard.

- Efficient algorithms for *hD* from $\mathbb{R}^d$ can be adapted to *ahD*.
- Fast C++ implementation for the R package `ddalpha` is in preparation, also for $\mathbb{S}^{d-1}$ with $d > 3$.

Applications to data analysis and further challenges:

- Visualisation, classification, spherical boxplots;
- Additional theory of *ahD*;
- Halfspace depth on manifolds other than $\mathbb{S}^{d-1}$?
  (Carrizosa, 1996; Dai and López-Pintado, 2021)

Selected Literature

- A. Ahidar-Coutrix and P. Berthet, Convergence of multivariate quantile surfaces, arXiv preprint arXiv:1607.02604, (2016).
- R. Fraiman and G. Muniz, Trimmed means for functional data, Test, 10 (2001), pp. 419–440.
- R. Y. Liu and K. Singh, Ordering directional data: concepts of data depth on circles and spheres, Ann. Statist., 20 (1992), pp. 1468–1484.
- S. Nagy and J. Dvořák, Illumination depth, J. Comput. Graph. Statist., 30 (2021), pp. 78–90.
- S. Nagy, S. Helander, G. Van Bever, L. Viitasaari, and P. Ilmonen, Flexible integrated functional depths, Bernoulli, 27 (2021), pp. 673–701.
- O. Pokotylo, P. Mozharovskyi, and R. Dyckerhoff, Depth and depth-based classification with R package ddalpha, Journal of Statistical Software, 91 (2019), pp. 1–46.
- C. G. Small, Measures of centrality for multivariate and directional distributions, Canad. J. Statist., 15 (1987), pp. 31–39.
- E. M. Werner, Illumination bodies and affine surface area, Studia Math., 110 (1994), pp. 257–269.

Conclusions: Depth and nonparametrics

The depth

- can extend nonparametrics to multivariate data;
- looks easy, but is (often) not;
- provides plenty of interesting open problems.

My conclusions:

- Read and talk to people. Especially outside your field.
- Not all is trivial. Even little progress counts.
- Think more, simulate less.

If you are interested, let us know

nagy@karlin.mff.cuni.cz

GeMS.karlin.mff.cuni.cz