An Introduction to Statistical Depth Functions

Satvik Saha
Supervised by Dr. Anirvan Chakraborty

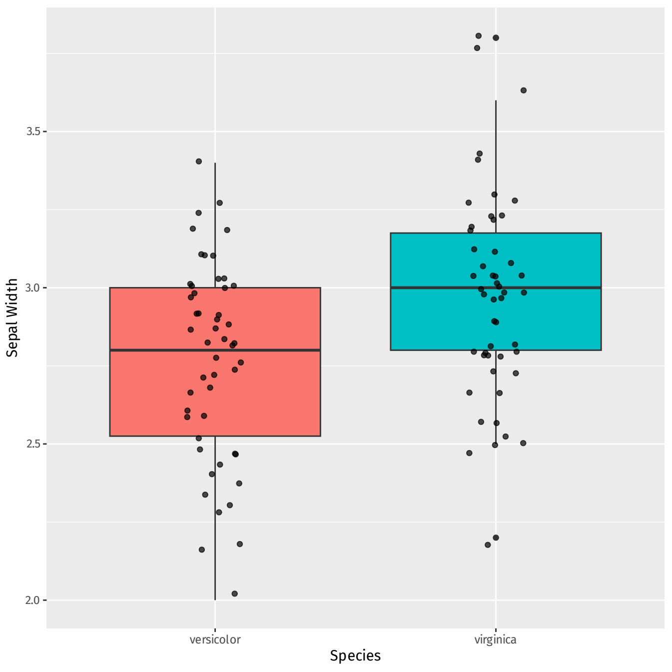
12 December, 2023

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

Outline

1. A two-sample testing problem
2. Depth Functions
3. The Depth-Depth plot
4. Depth based classification
5. Depth functions for Functional Data
6. Future work

A two-sample testing problem



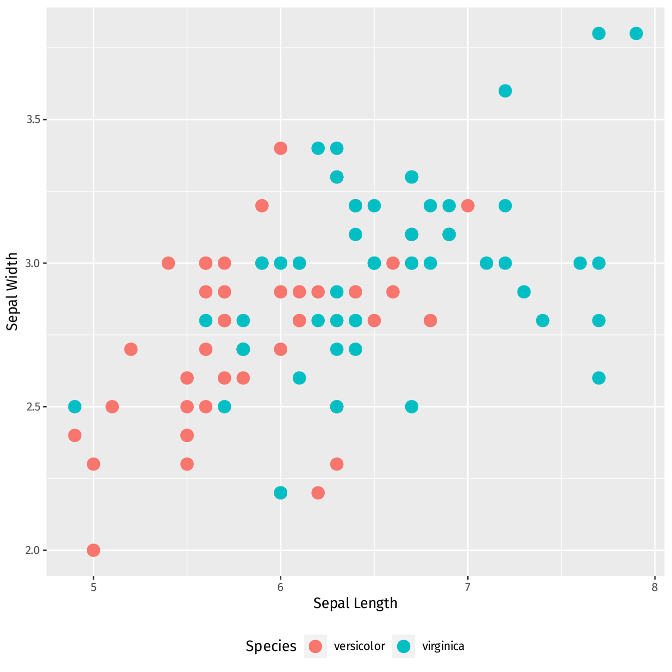
3.5
3.0
2.5
2.0
Sepal Width
versicolor
virginica
Species

Wilcoxon rank sum test

Given two random samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$, construct

$$W = \sum_j r(Y_j, \mathscr{D}_F \cup \mathscr{D}_G), \qquad r(Y, \mathscr{D}) = \sum_{Z \in \mathscr{D}} \mathbb{1}(Z \leq Y).$$

This is *distribution free* under the null hypothesis that both samples have the same underlying distribution.


Sepal Width
3.5
3.0
2.5
2.0
5
6
7
8
Sepal Length
Species
versicolor
virginica

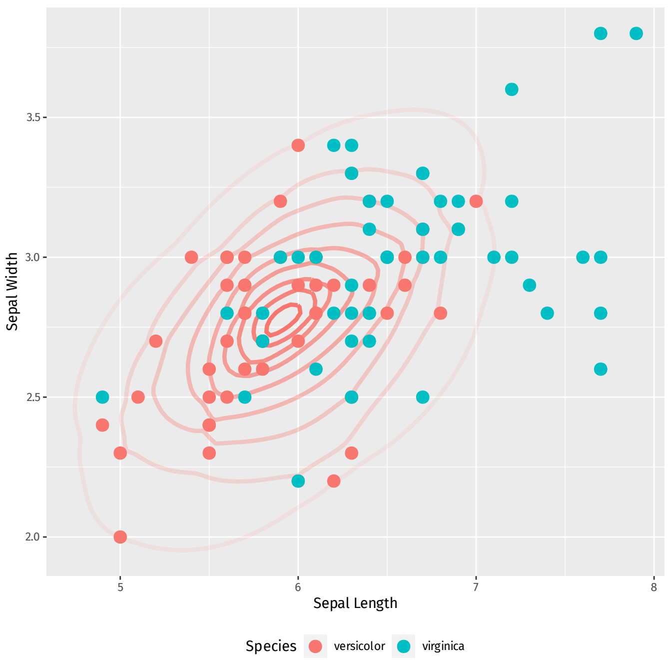
A generalization for multivariate data

Given *multivariate* data, we wish to construct

$$W^* = \sum_j r(Y_j, \mathscr{D}_F \cup \mathscr{D}_G), \qquad r^*(Y, \mathscr{D}) = \sum_{Z \in \mathscr{D}} 1(Z \text{ ?? } Y).$$

Furthermore, we want W^* to be able to detect differences in location and scale between F and G.

Liu, R.Y., & Singh, K. (1993) A Quality Index Based on Data Depth and Multivariate Rank Tests.


3.5
3.0
2.5
2.0
Sepal Width
5
6
7
8
Sepal Length
Species
versicolor
virginica

How do we quantify this notion of centrality?

Depth Functions

Depth Functions

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

Points which are *more central* are said to be *deeper*.

This framework allows many rank based nonparametric techniques to be translated to a broader class of data, e.g. multivariate and functional data.

Depth Functions

A *depth function* quantifies how central a point x is with respect to a distribution F.

Points which are *more central* are said to be *deeper*.

This framework allows many rank based nonparametric techniques to be translated to a broader class of data, e.g. multivariate and functional data.

Depth Functions

A *depth function* quantifies how central a point x is with respect to a distribution F.

Points which are *more central* are said to be *deeper*.

This framework allows many rank based nonparametric techniques to be translated to a broader class of data, e.g. multivariate and functional data.

Some applications of depth functions

1. Inference procedures
    - Hypothesis tests
    - Rank tests
    - Multivariate quantiles
    - Confidence regions
2. Exploratory data analysis
3. Classification and clustering
4. Outlier detection

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

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

1. **Affine invariance:** $D(A\mathbf{x} + b, F_{A\mathbf{x}+b}) = D(\mathbf{x}, F_X)$.
2. **Maximality at centre:** $D(\theta, F_X) = \sup_{\mathbf{x}\in\mathbb{R}^p} D(\mathbf{x}, F)$.
3. **Monotonicity along rays:** $D(\mathbf{x}, F) \leq D(\theta + \alpha(\mathbf{x} - \theta), F)$.
4. **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

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

1. **Affine invariance:** $D(A\mathbf{x} + b, F_{A\mathbf{x}+b}) = D(\mathbf{x}, F_X)$.
2. **Maximality at centre:** $D(\theta, F_X) = \sup_{\mathbf{x}\in\mathbb{R}^p} D(\mathbf{x}, F)$.
3. **Monotonicity along rays:** $D(\mathbf{x}, F) \leq D(\theta + \alpha(\mathbf{x} - \theta), F)$.
4. **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

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

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

1. **Affine invariance:** $D(A\mathbf{x} + b, F_{A\mathbf{x}+b}) = D(\mathbf{x}, F_X)$.
2. **Maximality at centre:** $D(\theta, F_X) = \sup_{\mathbf{x}\in\mathbb{R}^p} D(\mathbf{x}, F)$.
3. **Monotonicity along rays:** $D(\mathbf{x}, F) \leq D(\theta + \alpha(\mathbf{x} - \theta), F)$.
4. **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

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

1. **Affine invariance:** $D(A\boldsymbol{x} + b, F_{A\boldsymbol{x}+b}) = D(\boldsymbol{x}, F_X)$.
2. **Maximality at centre:** $D(\theta, F_X) = \sup_{\boldsymbol{x}\in\mathbb{R}^p} D(\boldsymbol{x}, F)$.
3. **Monotonicity along rays:** $D(\boldsymbol{x}, F) \leq D(\theta + \alpha(\boldsymbol{x} - \theta), F)$.
4. **Vanish at infinity:** $D(\boldsymbol{x}, F) \to 0$ as $\|\boldsymbol{x}\| \to \infty$.

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

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

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

1. **Affine invariance:** $D(A\mathbf{x} + b, F_{A\mathbf{x}+b}) = D(\mathbf{x}, F_X)$.
2. **Maximality at centre:** $D(\theta, F_X) = \sup_{\mathbf{x}\in\mathbb{R}^p} D(\mathbf{x}, F)$.
3. **Monotonicity along rays:** $D(\mathbf{x}, F) \leq D(\theta + \alpha(\mathbf{x} - \theta), F)$.
4. **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

Depth contours

The *region of depth d* is defined by

$$\mathcal{R}(d, F) = \{\mathbf{x} \in \mathbb{R}^p \mid D(\mathbf{x}, F) \geq d\}.$$

The boundary $\partial\mathcal{R}(d, F)$ is called the *contour of depth d*.

Define

$$R(\mathbf{x}, F) = P(D(Y, F) \geq D(\mathbf{x}, F) \mid Y \sim F).$$

Then, as long as $D(\cdot, F)$ is continuous, the probability integral transform gives

$$R(X, F) \sim \text{Uniform}[0, 1].$$

Liu, R.Y., Parelius, J.M, & Singh, K. (1999) Multivariate analysis by data depth: descriptive statistics, graphics and inference

Depth contours

The *region of depth d* is defined by

$$\mathcal{R}(d, F) = \{\mathbf{x} \in \mathbb{R}^p \mid D(\mathbf{x}, F) \geq d\}.$$

The boundary $\partial\mathcal{R}(d, F)$ is called the *contour of depth d*.

Define

$$R(\mathbf{x}, F) = P(D(Y, F) \geq D(\mathbf{x}, F) \mid Y \sim F).$$

Then, as long as $D(\cdot, F)$ is continuous, the probability integral transform gives

$$R(X, F) \sim \text{Uniform}[0, 1].$$

Liu, R.Y., Parelius, J.M, & Singh, K. (1999) Multivariate analysis by data depth: descriptive statistics, graphics and inference

Depth contours

The *region of depth d* is defined by

$$\mathcal{R}(d, F) = \{\mathbf{x} \in \mathbb{R}^p \mid D(\mathbf{x}, F) \geq d\}.$$

The boundary $\partial\mathcal{R}(d, F)$ is called the *contour of depth d*.

Define

$$R(\mathbf{x}, F) = P(D(Y, F) \geq D(\mathbf{x}, F) \mid Y \sim F).$$

Then, as long as $D(\cdot, F)$ is continuous, the probability integral transform gives

$$R(X, F) \sim \text{Uniform}[0, 1].$$

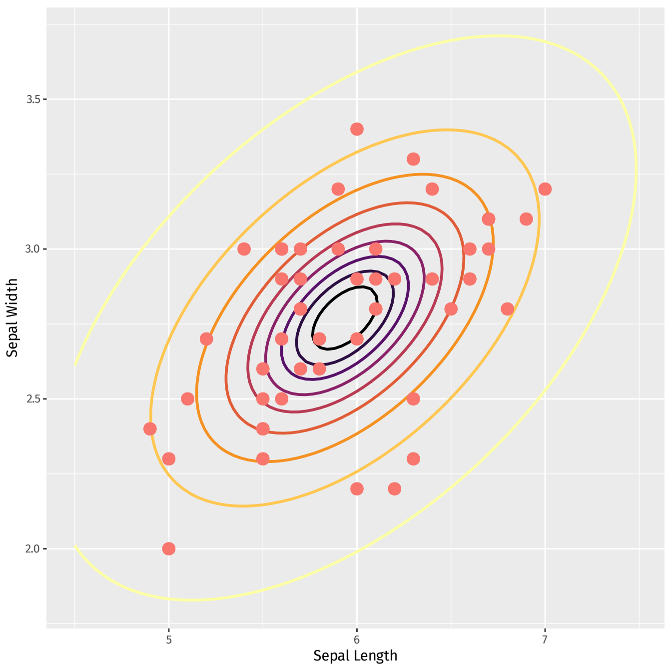
Liu, R.Y., Parelius, J.M, & Singh, K. (1999) Multivariate analysis by data depth: descriptive statistics, graphics and inference

Mahalanobis depth

Produces elliptic contours, using the first two moments of the given distribution.

$$D_{Mh}(x, F) = \frac{1}{1 + (x - \mu)\Sigma^{-1}(x - \mu)}.$$

A robust version can be obtained by using MCD estimators.

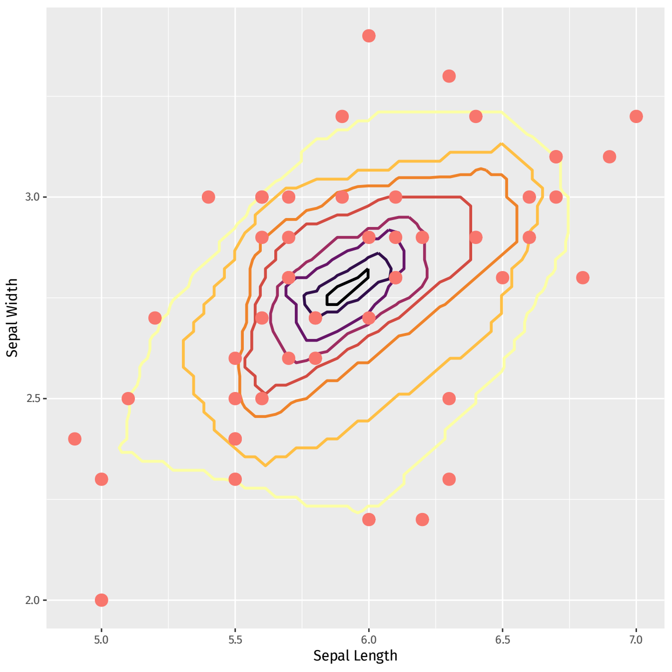


3.5
3.0
2.5
2.0
Sepal Width
5
6
7
Sepal Length

Halfspace/Tukey depth

Given a point $\boldsymbol{x} \in \mathbb{R}^p$, examine all hyperplanes through $\boldsymbol{x}$, and find the halfspace with the least probability.

$$D_H(x, F) = \inf_{v \in \mathbb{R}^p \setminus \{0\}} P(\underbrace{v^\top X \leq v^\top x}_{\substack{X \text{ is in a halfspace} \\ \text{through } x}}).$$



Sepal Width
3.0
2.5
2.0
5.0
5.5
6.0
6.5
7.0
Sepal Length

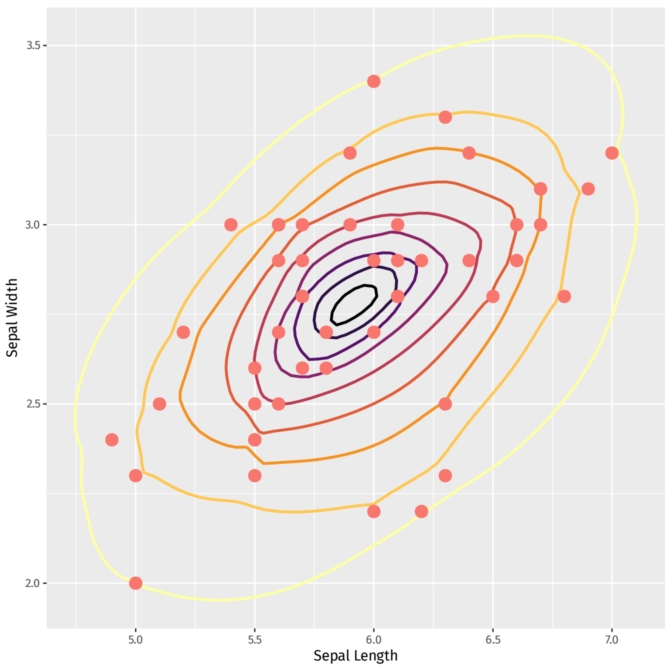
Spatial depth

Examine the average of unit vectors pointing out of $\boldsymbol{x}$.

$$D_{Sp}(x, F) = 1 - \left\| E\left[\underbrace{\frac{X - x}{\|X - x\|}}_{\substack{\text{unit vector} \\ \text{from } x \text{ to } X}} \right] \right\|.$$

Spatial depth is *not* always monotonic with respect to the deepest point.

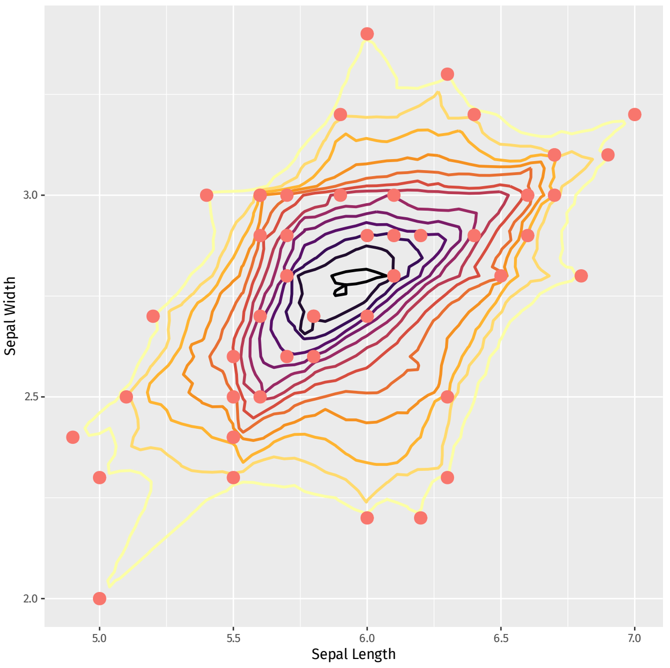
Nagy., S. (2017) Monotonicity properties of spatial depth


3.5
3.0
2.5
2.0
Sepal Width
5.0
5.5
6.0
6.5
7.0
Sepal Length

Simplicial depth

Examine the probability of $\boldsymbol{x}$ being contained in a random simplex.

$$D_S(x, F) = P(x \in \text{simplex}[X_1, \ldots, X_{p+1}] \mid X_i \overset{iid}{\sim} F).$$


Sepal Width
Sepal Length
3.0
2.5
2.0
5.0
5.5
6.0
6.5
7.0

Projection depth

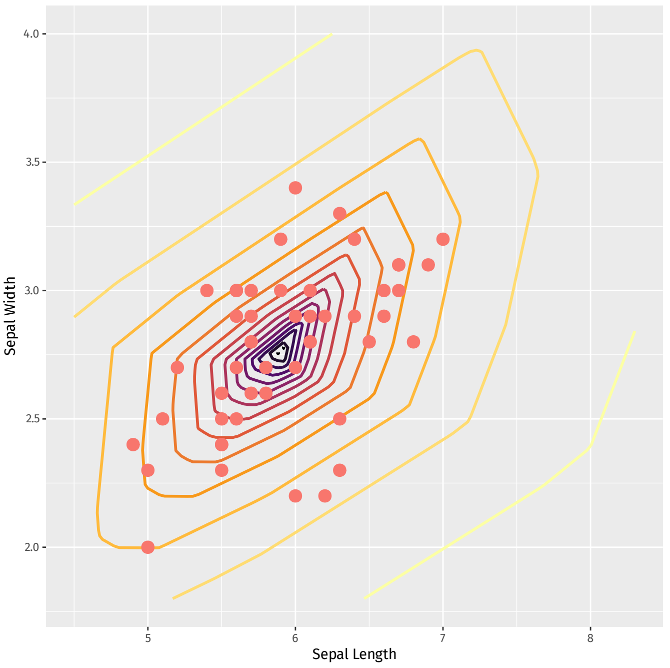
Examine the maximum outlyingness of $\boldsymbol{x}$ with respect to projections.

$$D_P(x, F) = \left(1 + \sup_{\|v\|=1} \frac{v^\top x - \mu(v^\top X)}{\sigma(v^\top X)}\right)^{-1}, \quad X \sim F.$$

A robust version can be defined as

$$D_P^*(x, F) = \left(1 + \sup_{\|v\|=1} \frac{v^\top x - \mathrm{median}(v^\top X)}{\mathrm{MAD}(v^\top X)}\right)^{-1}, \quad X \sim F,$$

$$\mathrm{MAD}(Y) = \mathrm{median}(|Y - \mathrm{median}(Y)|).$$



4.0
3.5
3.0
2.5
2.0
Sepal Width
5
6
7
8
Sepal Length

Why not use likelihood contours?

Why not use likelihood contours?

The ‘Curse of Dimensionality’.

The Depth-Depth plot

Depth-Depth plots

Let F, G be two distributions on $\mathbb{R}^p$, and let D be a depth function. We construct the D-D plot

$$DD(F, G) = \{(D(x, F),\ D(x, G)) : x \in \mathbb{R}^p\}.$$

Given data $\mathscr{D}_F, \mathscr{D}_G$, we may instead look at the D-D plot

$$DD(\hat{F}_m, \hat{G}_n) = \left\{ \left(D(x, \hat{F}_m),\ D(x, \hat{G}_n) \right) : x \in \mathscr{D}_F \cup \mathscr{D}_G \right\}.$$

Liu, R.Y., Parelius, J.M, & Singh, K. (1999) Multivariate analysis by data depth: descriptive statistics, graphics and inference

Depth-Depth plots

Let F, G be two distributions on $\mathbb{R}^p$, and let D be a depth function. We construct the D-D plot

$$DD(F, G) = \{(D(x, F),\ D(x, G)) : x \in \mathbb{R}^p\}.$$

Given data $\mathcal{D}_F, \mathcal{D}_G$, we may instead look at the D-D plot

$$DD(\hat{F}_m, \hat{G}_n) = \left\{\left(D(x, \hat{F}_m),\ D(x, \hat{G}_n)\right) : x \in \mathcal{D}_F \cup \mathcal{D}_G\right\}.$$

Liu, R.Y., Parelius, J.M, & Singh, K. (1999) Multivariate analysis by data depth: descriptive statistics, graphics and inference

Depth-Depth plots

Let F, G be two distributions on $\mathbb{R}^p$, and let D be a depth function. We construct the D-D plot

$$DD(F, G) = \{(D(\boldsymbol{x}, F),\ D(\boldsymbol{x}, G)) : \boldsymbol{x} \in \mathbb{R}^p\}.$$

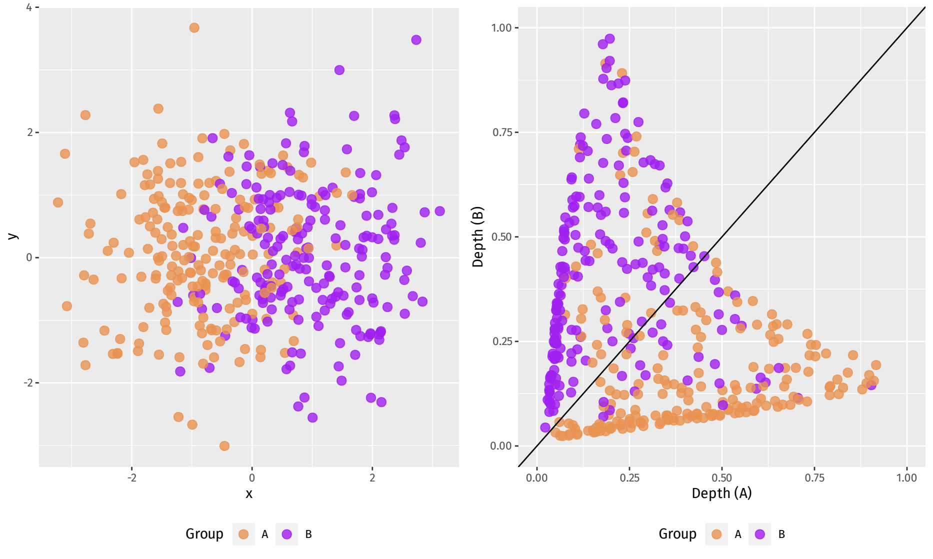
Given data $\mathcal{D}_F, \mathcal{D}_G$, we may instead look at the D-D plot

$$DD(\hat{F}_m, \hat{G}_n) = \left\{ \left(D(\boldsymbol{x}, \hat{F}_m),\ D(\boldsymbol{x}, \hat{G}_n) \right) : \boldsymbol{x} \in \mathcal{D}_F \cup \mathcal{D}_G \right\}.$$

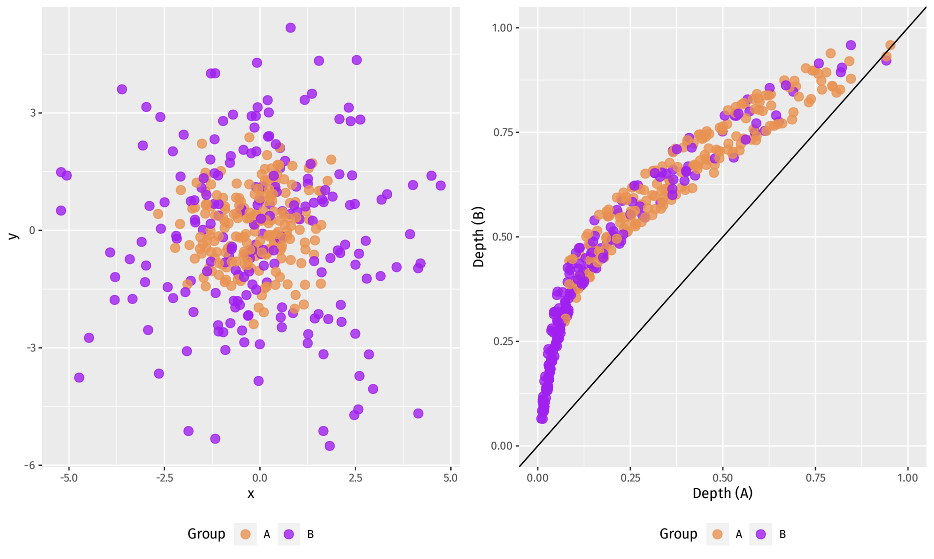
Liu, R.Y., Parelius, J.M, & Singh, K. (1999) Multivariate analysis by data depth: descriptive statistics, graphics and inference

Identical distributions

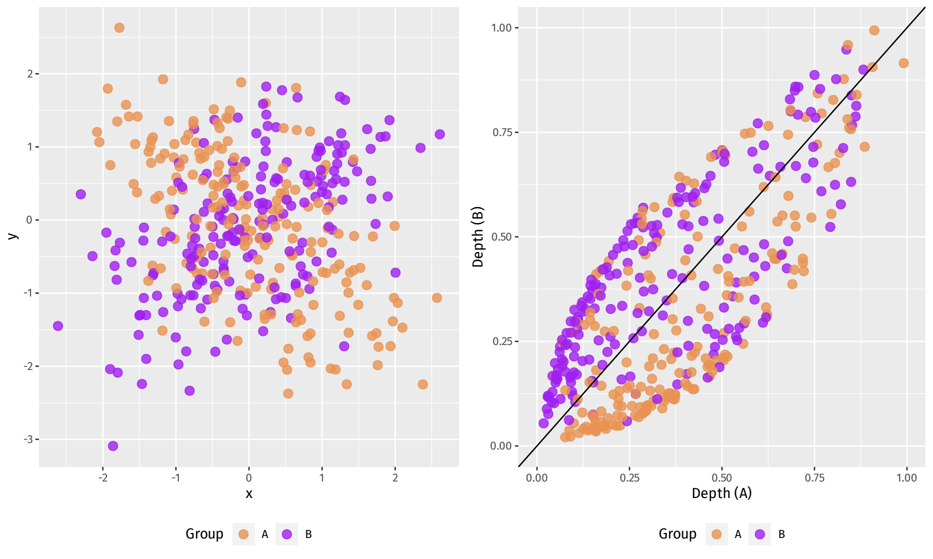
Location difference



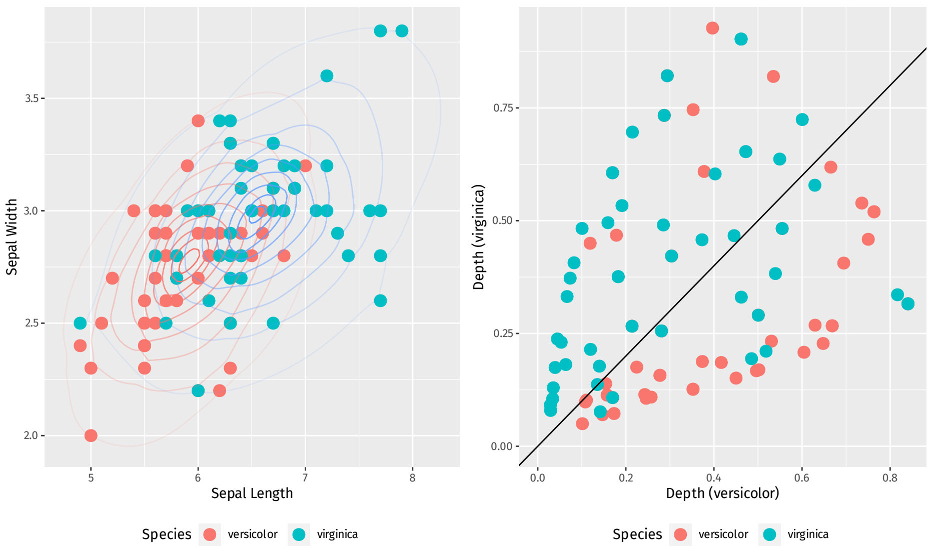
Scale difference



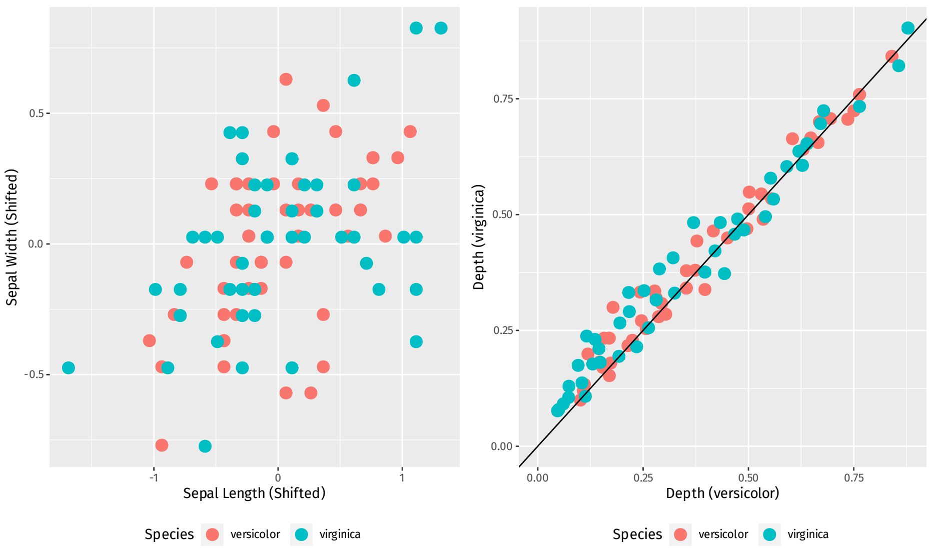
Scale difference



Location and scale difference



Sepal Width
3.5
3.0
2.5
2.0
5
6
7
8
Sepal Length
Species
versicolor
virginica
Depth (virginica)
0.75
0.50
0.25
0.00
0.0
0.2
0.4
0.6
0.8
Depth (versicolor)
Species
versicolor
virginica



Sepal Width (Shifted)
Sepal Length (Shifted)
Species
versicolor
virginica
Depth (virginica)
Depth (versicolor)
Species
versicolor
virginica

Depth based classification

Maximum depth classifiers

Given a point $\boldsymbol{x} \in \mathbb{R}^p$, assign it to the class with respect to which it has maximum depth. In other words, choose

$$\hat{k}(\boldsymbol{x}) = \arg\max_j D(\boldsymbol{x}, \hat{F}_j).$$

Under certain conditions, this asymptotically performs on par with the Bayes classifier.

Ghosh, A.K., & Chaudhuri, P. (2005) On maximum depth and related classifiers


Train
Test
Sepal Width
Sepal Length
4.0
3.5
3.0
2.5
2.0
5
6
7
8
Species
versicolor
virginica
Predicted species
versicolor
virginica

The *relative data depth*

$$\mathrm{ReD}(\boldsymbol{x}) = D(\boldsymbol{x}, \hat{F}_{\hat{k}(\boldsymbol{x})}) - \max_{j \neq \hat{k}(\boldsymbol{x})} D(\boldsymbol{x}, \hat{F}_j)$$

gives a measure of confidence in the classification of $\boldsymbol{x}$.

Jörnsten, R. (2004) Clustering and classification based on the L_1 data depth

5
0
-5
y
-5
0
5
x
Group
A
B
1.00
0.75
0.50
0.25
0.00
Depth (B)
0.00
0.25
0.50
0.75
1.00
Depth (A)
Group
A
B

Depth-Depth classifiers

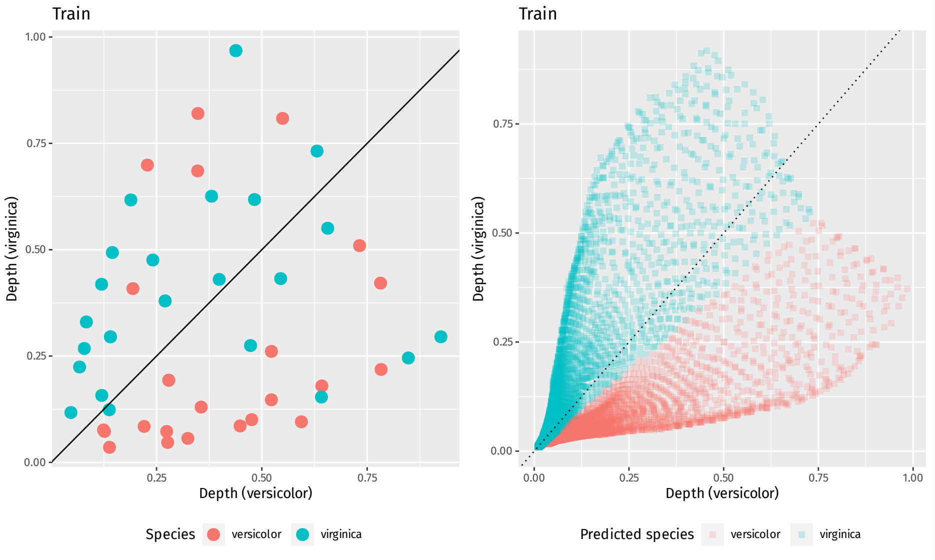
Given data $\mathscr{D}_F, \mathscr{D}_G$, look at the D-D plot

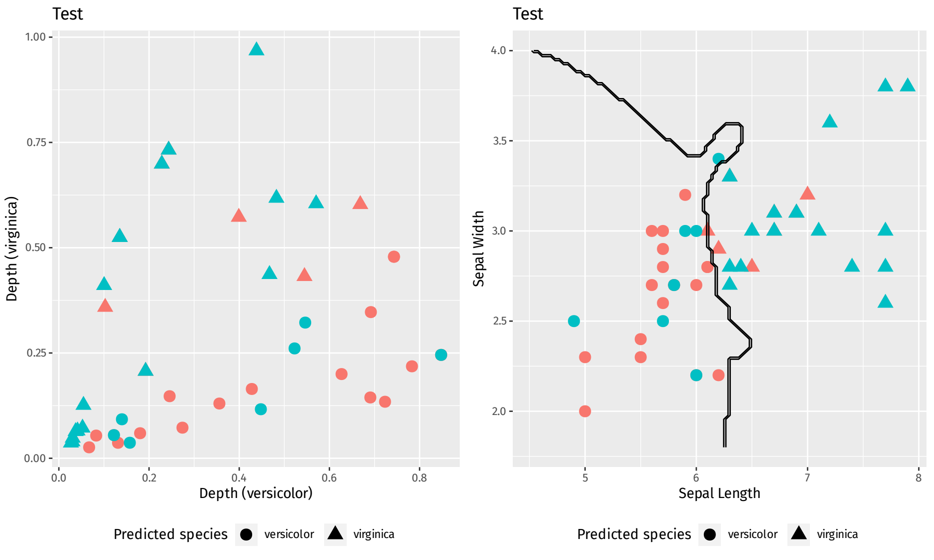
$$DD(\hat{F}_m, \hat{G}_n) = \left\{ \Big(D(\mathbf{x}_i, \hat{F}_m),\ D(\mathbf{x}_i, \hat{G}_n) \Big) : \mathbf{x}_i \in \mathscr{D}_F \cup \mathscr{D}_G \right\},$$

and find a function ϕ which separates points from the two classes.

For $\boldsymbol{x} \in \mathbb{R}^p$, check which region the point $(D(\boldsymbol{x}, \hat{F}_m),\ D(\boldsymbol{x}, \hat{G}_n))$ lies in, and assign it to the corresponding class.

Li, J., Cuestas-Albertos, J.A., & Liu, R.Y. (2012) DD-Classifier: Nonparametric Classification Procedure Based on DD-Plot


Train
1.00
0.75
0.50
0.25
0.00
Depth (virginica)
0.25
0.50
0.75
Depth (versicolor)
Species
versicolor
virginica
Train
0.75
0.50
0.25
0.00
Depth (virginica)
0.00
0.25
0.50
0.75
1.00
Depth (versicolor)
Predicted species
versicolor
virginica


Test
1.00
0.75
0.50
0.25
0.00
Depth (virginica)
0.0
0.2
0.4
0.6
0.8
Depth (versicolor)
Predicted species
versicolor
virginica
Test
4.0
3.5
3.0
2.5
2.0
Sepal Width
5
6
7
8
Sepal Length
Predicted species
versicolor
virginica

Elliptic distributions

Suppose that the underlying population distributions are elliptic, i.e. their density functions are of the form

$$c_i|\Sigma_i|^{-1/2}\, h_i\left((x-\mu_i)^\top \Sigma_i^{-1}(x-\mu_i)\right)$$

for strictly decreasing functions h_i. Then, the *Mahalanobis*, *simplicial*, and *projection* depths $D(\cdot, F_i)$ are strictly increasing functions of the respective densities.

Thus, the Bayes rule involves comparing $\phi(D(x,F))$ and $D(x,G)$ for some strictly increasing function ϕ.

Li, J., Cuestas-Albertos, J.A., & Liu, R.Y. (2012) DD-Classifier: Nonparametric Classification Procedure Based on DD-Plot

Depth functions for Functional Data

Integrated, infimal, and random projection depths

$$D_{int}(X, F_X) = \int_T D(X(t), F_{X(t)})\, w(t)\, dt.$$

$$D_{inf}(X, F_X) = \inf_{t \in T} D(X(t), F_{X(t)}).$$

$$D_{RP}(X, F_X) = \inf_{\phi} D(\langle X, \phi \rangle, F_{\langle X, \phi \rangle}).$$

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

Outlyingness matrices

Given a random p-variate function X, define a pointwise outlyingness function as

$$\boldsymbol{O}(X(t), F_{X(t)}) = \left[\frac{1}{D(X(t), F_{X(t)})} - 1\right] \cdot \boldsymbol{v}(t).$$

With this, define

$$\boldsymbol{MO}(X, F_X) = \int_T \boldsymbol{O}(X(t), F_{X(t)})\, w(t)\, dt,$$

$$VO(X, F_X) = \int_T \|\boldsymbol{O}(X(t), F_{X(t)}) - \boldsymbol{MO}(X, F_X)\|^2\, w(t)\, dt.$$

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

Outlyingness matrices

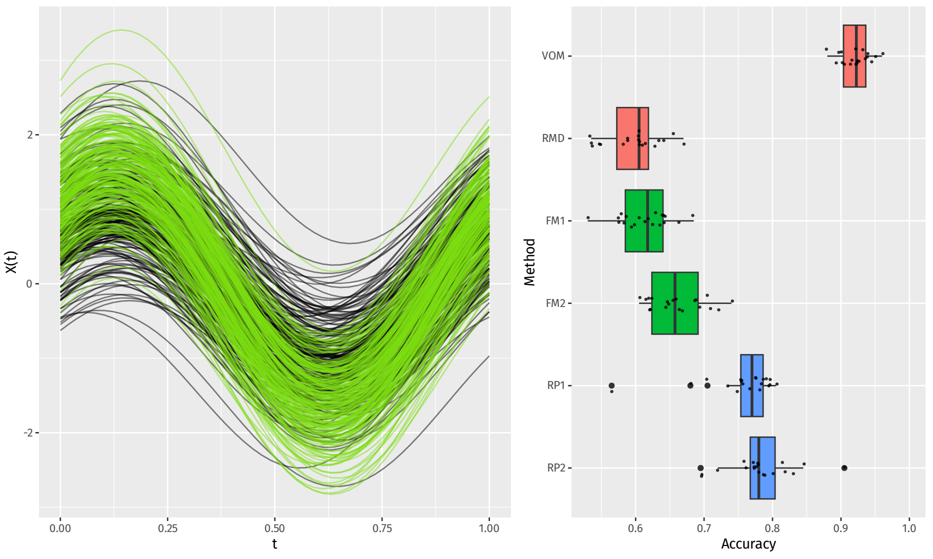
Furthermore, denoting

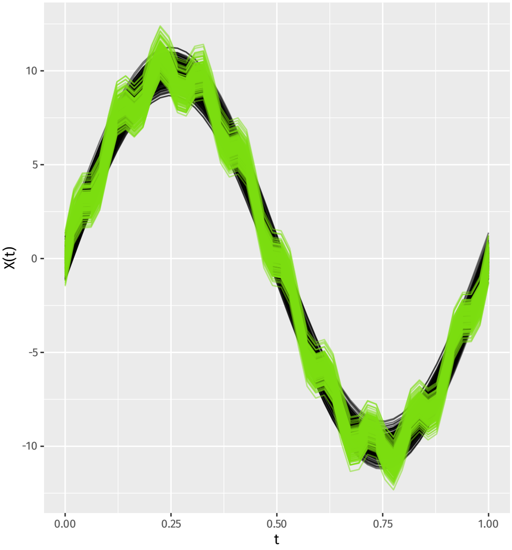
$$\tilde{\boldsymbol{O}}(X(t), F_{X(t)}) = \boldsymbol{O}(X(t), F_{X(t)}) - \boldsymbol{MO}(X, F_X),$$

define the *variational outlyingness matrix*

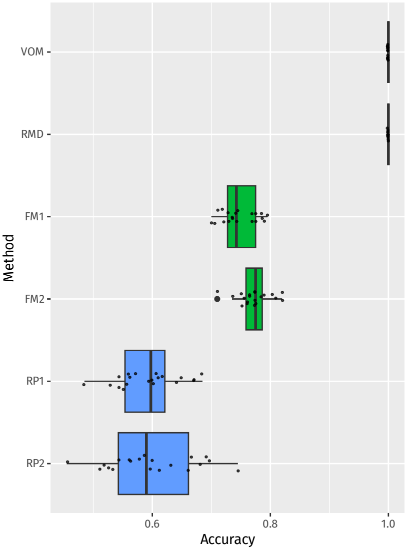
$$VOM(X, F_X) = \int_T \tilde{\boldsymbol{O}}(X(t), F_{X(t)})\, \tilde{\boldsymbol{O}}(X(t), F_{X(t)})^\top\, w(t)\, dt.$$

Use either the feature vector $(\boldsymbol{MO}^\top, VO)^\top$ or $\|VOM\|$ for classification.

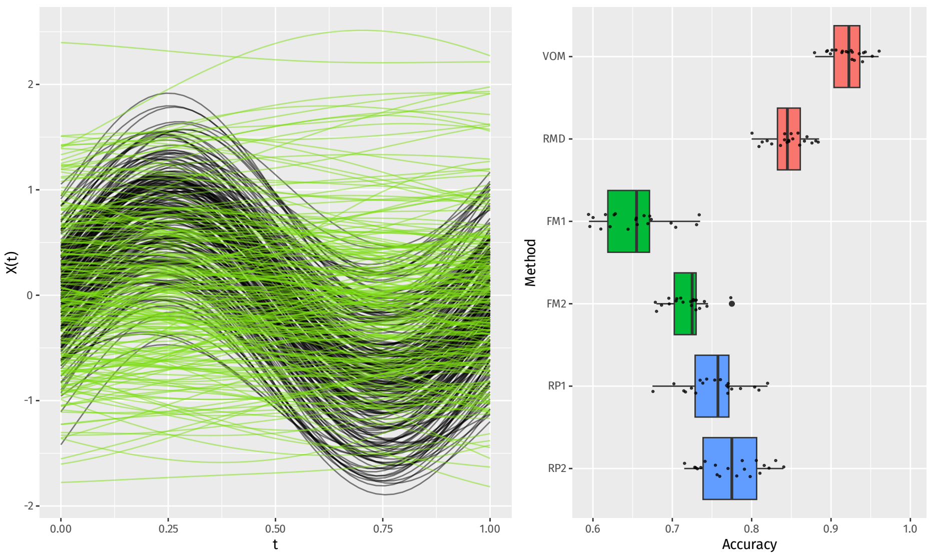

X(t)
2
0
-2
0.00
0.25
0.50
0.75
1.00
t
Method
VOM
RMD
FM1
FM2
RP1
RP2
0.6
0.7
0.8
0.9
1.0
Accuracy



X(t)
10
5
0
-5
-10
0.00
0.25
0.50
0.75
1.00
t

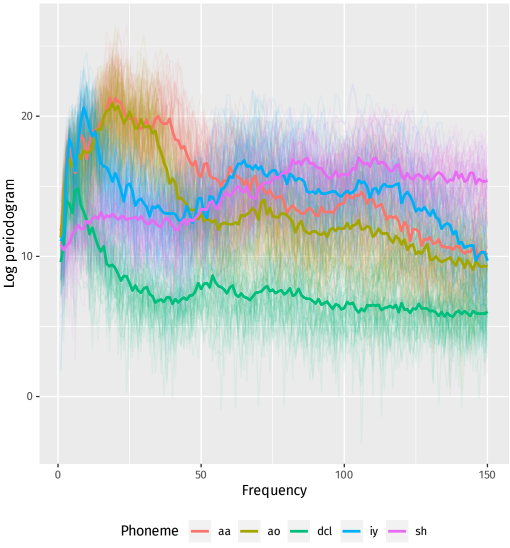


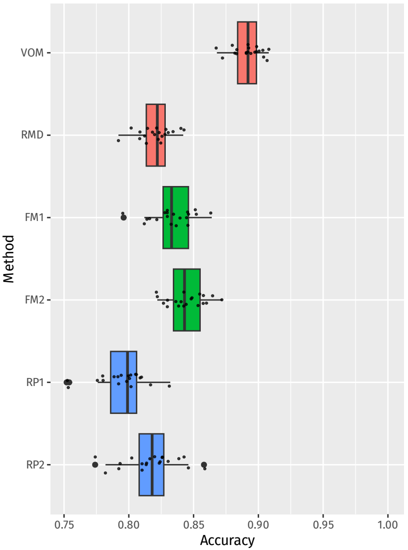
Method
VOM
RMD
FM1
FM2
RP1
RP2
0.6
0.8
1.0
Accuracy



X(t)
2
1
0
-1
-2
0.00
0.25
0.50
0.75
1.00
t
Method
VOM
RMD
FM1
FM2
RP1
RP2
0.6
0.7
0.8
0.9
1.0
Accuracy

Phonemes in digitized speech



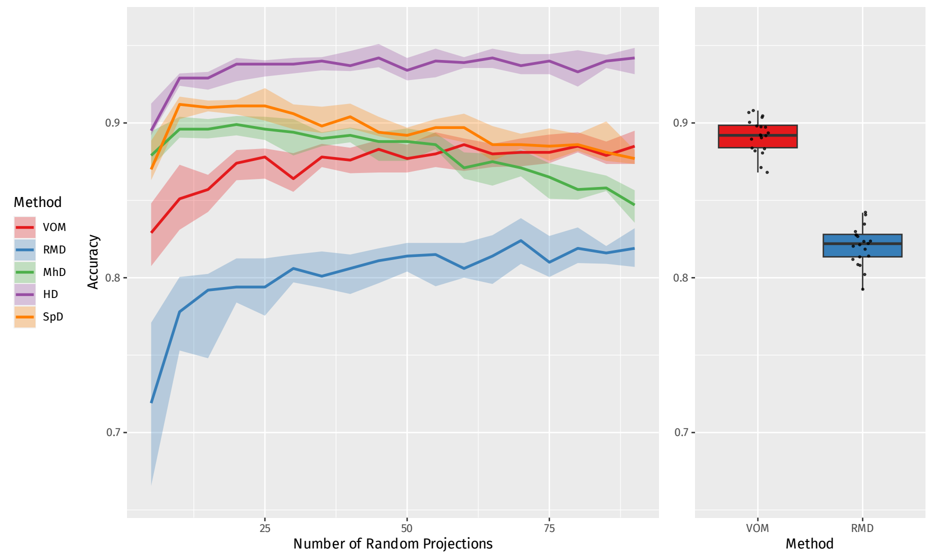


Functional $\to$ Multivariate, via random projections

Replace $\{X(t)\}_{t \in T}$ with $\{\langle X, \phi_j \rangle\}_{j=1}^{\ell}$, where $\phi_1, \ldots, \phi_\ell$ are random functions and

$$\langle X, \phi \rangle = \int_T \langle X(t), \phi(t) \rangle \, w(t) \, dt.$$

Phonemes in digitized speech revisited



Do depth functions completely characterize probability distributions?

Do depth functions completely characterize probability distributions?

Sometimes!

Halfspace depth revisited

The halfspace depth characterizes discrete probability distributions, i.e. if $D_H(\cdot, P) = D_H(\cdot, Q)$ and one of P, Q is discrete, then $P = Q$.

The halfspace depth also characterizes elliptic probability distributions.

Cuesta-Albertos, J.A., & Nieto-Reyes, A. (2008) The Tukey and the random Tukey depths characterize discrete distributions

Kong, L., & Zuo, Y. (2010) Smooth depth contours characterize the underlying distribution

A counterexample

Consider $X \sim P$, $Y \sim Q$ where

$$\psi_X(\boldsymbol{t}) = \exp(-\|\boldsymbol{t}\|_1^{1/2}), \qquad \psi_Y(\boldsymbol{t}) = \exp(-\|\boldsymbol{t}\|_{1/2}^{1/2}).$$

Observe that the *marginals* of X and Y are identically distributed!

This is because they have the same characteristic function,

$$\psi(t) = \exp(-|t|^{1/2}).$$

Nagy, S. (2021) Halfspace depth does not characterize probability distributions

A counterexample

Next, if $\psi_Z(\boldsymbol{t}) = \psi(\|\boldsymbol{t}\|_\alpha)$, then $\boldsymbol{v}^\top \boldsymbol{Z} \stackrel{d}{=} \|\boldsymbol{v}\|_\alpha Z_1$. Such distributions are called α-*symmetric*.

Using this, it can be shown that

$$D_H(\boldsymbol{x}, P) = D_H(\boldsymbol{x}, Q) = F(-\|\boldsymbol{x}\|_\infty),$$

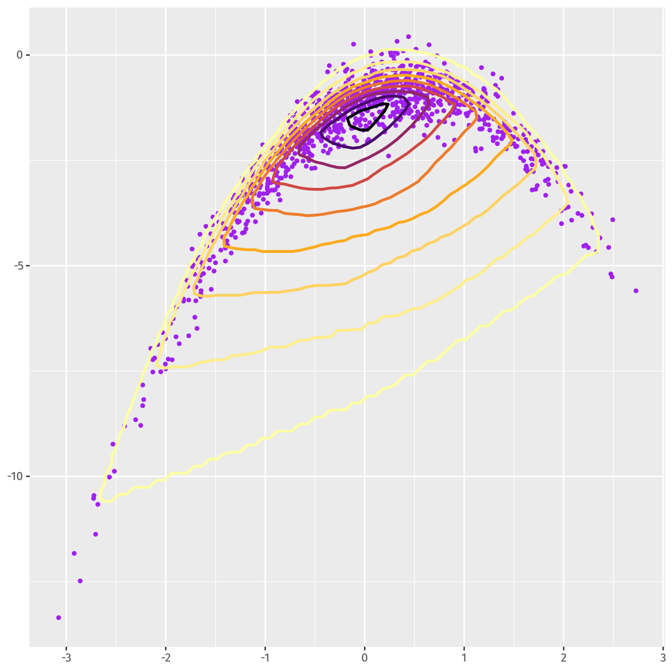
where F is the cdf of X_1.

Future work

Local depths

The notions of depth discussed so far work well with elliptic, unimodal distributions, but fail to capture the natures of more general distributions.

Agostinelli, C., & Romanazzi, M. (2011) Local depth


0
-5
-10
-3
-2
-1
0
1
2
3

Distribution-free procedures

Use ideas from optimal transportation to investigate more canonical notions of depth (for instance, the Monge-Kantorovich depth), and thereby establish procedures independent of the underlying distributions/spaces.

Chernozhukov, V., Galichon, A., Hallin, M., & Henry, M. (2017) Monge–Kantorovich depth, quantiles, ranks and signs

References i

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

[2] 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.

[3] 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.

[4] J. A. 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.

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

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

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

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

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

[10] 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.

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

[12] Jun Li, Juan A. 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.

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

[14] 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.

[15] 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.

References iv

[16] 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.

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

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

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

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

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

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

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