STAT 593
Robust statistics:
Depth and robust estimators

Joseph Salmon
http://josephsalmon.eu

Télécom Paristech, Institut Mines-Télécom
&
University of Washington, Department of Statistics
(Visiting Assistant Professor)

Outline

Table of Contents

1D case ($p = 1$)

Here, $X = [x_1, \ldots, x_n]$ is 1-dimensional : $\forall i \in [n], x_i \in \mathbb{R}$

Definition

For a fixed dataset X and for any point $x \in \mathbb{R}$, we defined the **depth**[1] of x w.r.t. X as

$$\text{depth}_1(x, X) = \min(\#\{i \in [n] : x_i \leq x\}, \#\{i \in [n] : x_i \geq x\})$$

Interpretation: the depth of a point in a dataset X is the minimum number of data points x_i on the left and on the right of x.

- $\text{depth}_1(x, X) = 0$ if $x > \max\limits_{i=1,\ldots,n}(x_i)$ or $x < \min\limits_{i=1,\ldots,n}(x_i)$
- $\text{depth}_1(x, X) = 1$ if[2] $x = \max\limits_{i=1,\ldots,n}(x_i)$ or $x = \min\limits_{i=1,\ldots,n}(x_i)$
- $\text{depth}_1(\text{Med}_n(x), X) \approx \frac{n}{2}$

[1] J. W. Tukey. "Mathematics and the picturing of data". In: *Proceedings of the International Congress of Mathematicians, Vancouver, 1975*. Vol. 2. 1975, pp. 523–531.

[2] when extrema are reached by only one point

Generalization in higher dimension

<u>Notation</u>: for a dataset X and a vector u, $\langle u, X\rangle$ is the dataset:

$$\langle u, X\rangle = [\langle u, x_1\rangle, \ldots, \langle u, x_n\rangle]$$

Definition

For a fixed dataset X and for any point $x \in \mathbb{R}^o$, we defined the depth of x w.r.t. X as

$$\mathrm{depth}_p(x, X) = \min_{\|u\|=1} \mathrm{depth}_1(\langle u, x\rangle, \langle u, X\rangle)$$

Generalization in higher dimension

Notation: for a dataset X and a vector u, $\langle u, X\rangle$ is the dataset:

$$\langle u, X\rangle = [\langle u, x_1\rangle, \ldots, \langle u, x_n\rangle]$$

Definition

For a fixed dataset X and for any point $x \in \mathbb{R}^o$, we defined the depth of x w.r.t. X as

$$\begin{aligned} \mathrm{depth}_p(x, X) &= \min_{\|u\|=1} \mathrm{depth}_1(\langle u, x\rangle, \langle u, X\rangle) \\ &= \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i\rangle \leq \langle u, x\rangle\}) \end{aligned}$$

Generalization in higher dimension

Notation: for a dataset X and a vector u, $\langle u, X\rangle$ is the dataset:

$$\langle u, X\rangle = [\langle u, x_1\rangle, \ldots, \langle u, x_n\rangle]$$

Definition

For a fixed dataset X and for any point $x \in \mathbb{R}^o$, we defined the depth of x w.r.t. X as

$$\begin{aligned}
\mathrm{depth}_p(x, X) &= \min_{\|u\|=1} \mathrm{depth}_1(\langle u, x\rangle, \langle u, X\rangle) \\
&= \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i\rangle \leq \langle u, x\rangle\}) \\
&= \min_{\|u\|=1} \#\{i \in [n] : x_i \in H_{u,x}\})
\end{aligned}$$

where we write $H_{u,x} = \{y \in \mathbb{R}^p : \langle u, y\rangle \leq \langle u, x\rangle\}$ for the half-space parametrized by a point x and a direction u, with $\|u\| = 1$

Generalization in higher dimension

Notation: for a dataset X and a vector u, $\langle u, X\rangle$ is the dataset:

$$\langle u, X\rangle = [\langle u, x_1\rangle, \ldots, \langle u, x_n\rangle]$$

Definition

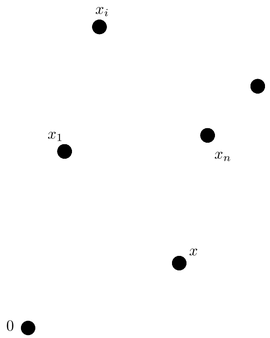
For a fixed dataset X and for any point $x \in \mathbb{R}^o$, we defined the depth of x w.r.t. X as

$$\begin{aligned}
\mathrm{depth}_p(x, X) &= \min_{\|u\|=1} \mathrm{depth}_1(\langle u, x\rangle, \langle u, X\rangle) \\
&= \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i\rangle \leq \langle u, x\rangle\}) \\
&= \min_{\|u\|=1} \#\{i \in [n] : x_i \in H_{u,x}\})
\end{aligned}$$

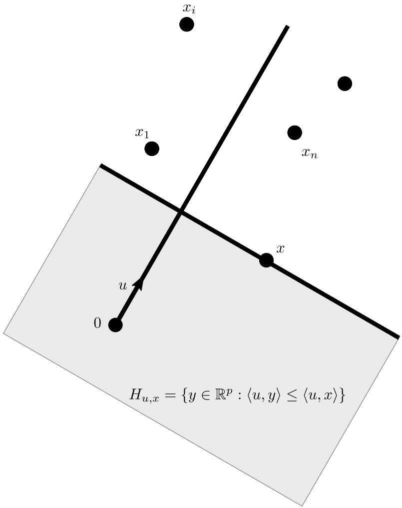
where we write $H_{u,x} = \{y \in \mathbb{R}^p : \langle u, y\rangle \leq \langle u, x\rangle\}$ for the half-space parametrized by a point x and a direction u, with $\|u\| = 1$

Interpretation: this is the least depth of x after any projection of the dataset on a space of dimension 1.

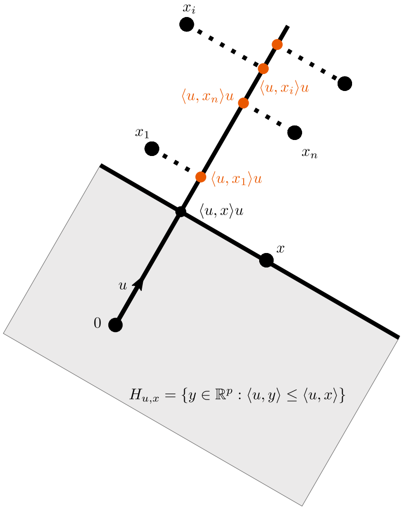
Visualization:



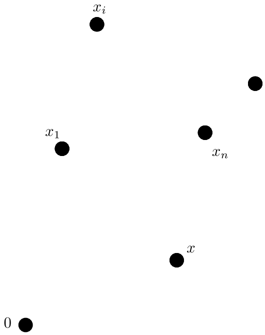
Visualization:



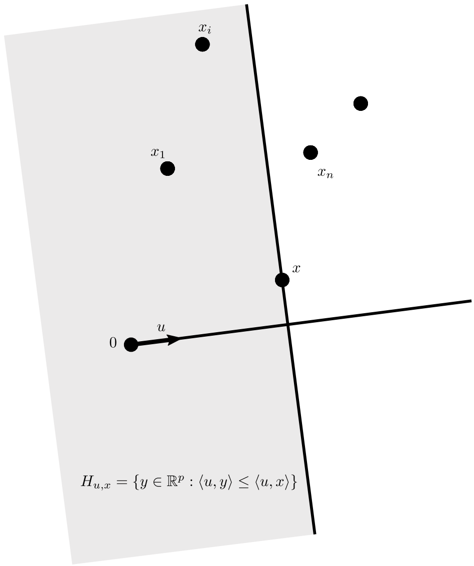
Visualization: $\mathrm{depth}_1(\langle u, x\rangle, \langle u, X\rangle) = 0$



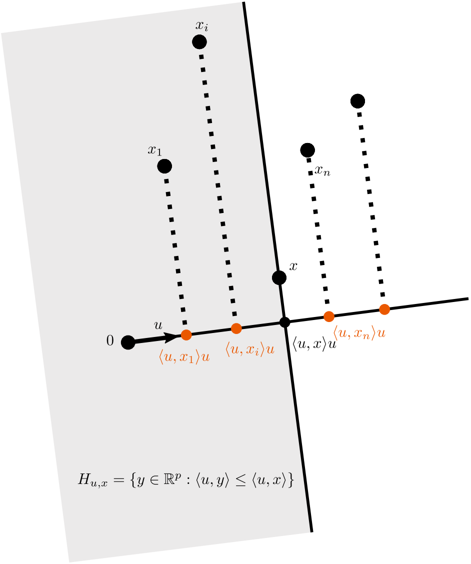
Visualization:



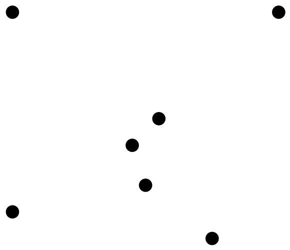
Visualization:



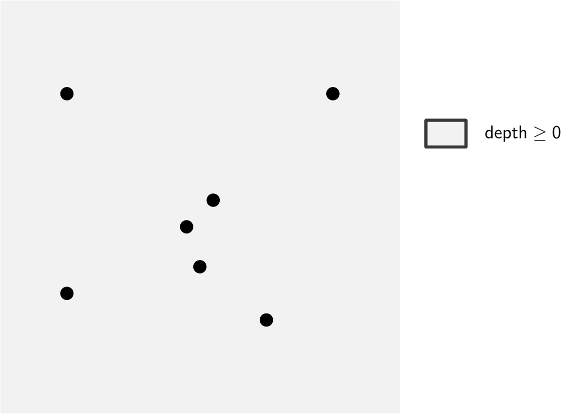
Visualization: $\mathrm{depth}_1(\langle u, x\rangle, \langle u, X\rangle) = 2$



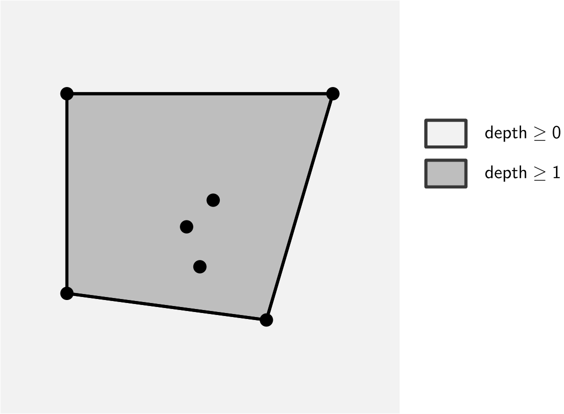
Example of depth values



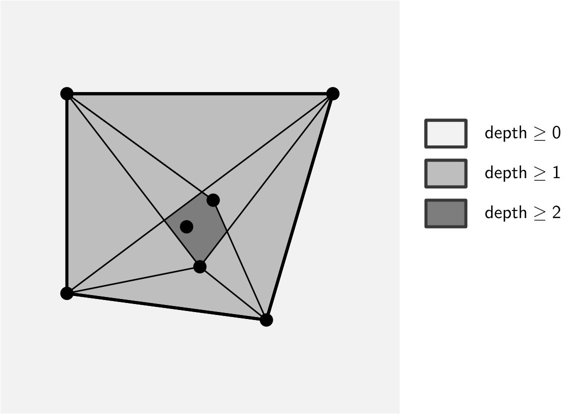
Example of depth values



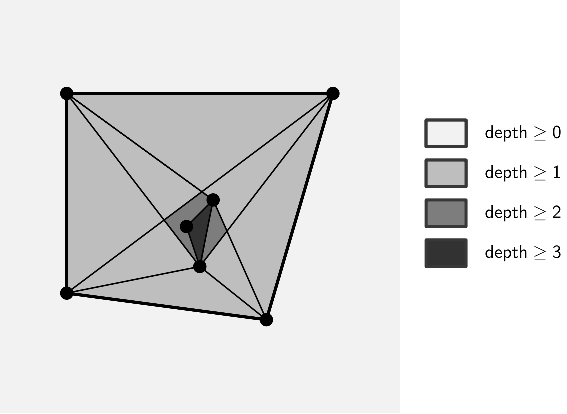
Example of depth values



Example of depth values



Example of depth values



Super-level set depth

Proposition

For any dataset X the super-level set of $x \mapsto \mathrm{depth}_p(x, X)$, *i.e.*, the sets $\{x \in \mathbb{R}^p : \mathrm{depth}_p(x, X) \geq t\}$ for any $t \geq 0$, are convex.

Proof: see Donoho and Gasko (1992)

Proposition

The set $\{x \in \mathbb{R}^p : \mathrm{depth}_p(x, X) \geq 1\}$ is the convex hull of the points $x_1, \ldots, x_n$.

Growth

Proposition

For any dataset X and Y, and any point x:

$$\mathrm{depth}_p(x, X) \leq \mathrm{depth}_p(x, X \cup Y)$$

Interpretation: the depth is non-decreasing w.r.t. merging datasets.

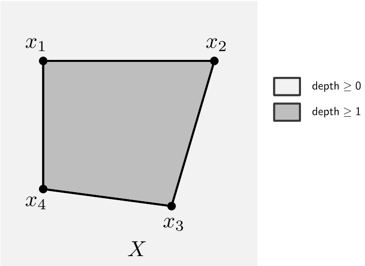
Growth

Proposition

For any dataset X and Y, and any point x:

$$\text{depth}_p(x, X) \leq \text{depth}_p(x, X \cup Y)$$

Interpretation: the depth is non-decreasing w.r.t. merging datasets.



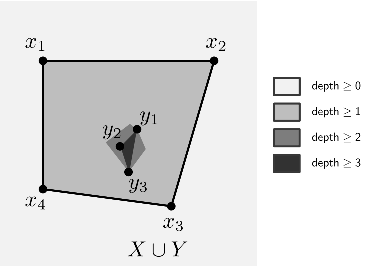
Growth

Proposition

For any dataset X and Y, and any point x:

$$\mathrm{depth}_p(x, X) \leq \mathrm{depth}_p(x, X \cup Y)$$

Interpretation: the depth is non-decreasing w.r.t. merging datasets.



depth_p is affine invariant

Proposition

For any $\mu \in \mathbb{R}^p$ and any non-singular matrix Σ

$$\mathrm{depth}_p(\Sigma x + \mu, \Sigma X + \mu) = \mathrm{depth}_p(x, X)$$

depth_p is affine invariant

Proposition

For any $\mu \in \mathbb{R}^p$ and any non-singular matrix Σ

$$\mathrm{depth}_p(\Sigma x + \mu, \Sigma X + \mu) = \mathrm{depth}_p(x, X)$$

Proof of proposition:

Fact 1: $\mathrm{depth}_p(x + \mu, X + \mu) = \mathrm{depth}_p(x, X)$ for any μ

depth_p is affine invariant

Proposition

For any $\mu \in \mathbb{R}^p$ and any non-singular matrix Σ

$$\mathrm{depth}_p(\Sigma x + \mu, \Sigma X + \mu) = \mathrm{depth}_p(x, X)$$

Proof of proposition:

Fact 1: $\mathrm{depth}_p(x + \mu, X + \mu) = \mathrm{depth}_p(x, X)$ for any μ

Proof of Fact 1:

$$\mathrm{depth}_p(x + \mu, X + \mu) = \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i + \mu \rangle \leq \langle u, x + \mu \rangle\})$$

depth_p is affine invariant

Proposition

For any $\mu \in \mathbb{R}^p$ and any non-singular matrix Σ

$$\mathrm{depth}_p(\Sigma x + \mu, \Sigma X + \mu) = \mathrm{depth}_p(x, X)$$

Proof of proposition:

<u>Fact 1</u>: $\mathrm{depth}_p(x + \mu, X + \mu) = \mathrm{depth}_p(x, X)$ for any μ

Proof of Fact 1:

$$\begin{aligned}\mathrm{depth}_p(x + \mu, X + \mu) &= \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i + \mu\rangle \le \langle u, x + \mu\rangle\}) \\ &= \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i\rangle \le \langle u, x\rangle\})\end{aligned}$$

depth_p is affine invariant

Proposition

For any $\mu \in \mathbb{R}^p$ and any non-singular matrix Σ

$$\mathrm{depth}_p(\Sigma x + \mu, \Sigma X + \mu) = \mathrm{depth}_p(x, X)$$

Proof of proposition:

Fact 1: $\mathrm{depth}_p(x + \mu, X + \mu) = \mathrm{depth}_p(x, X)$ for any μ

Proof of Fact 1:

$$\begin{aligned}
\mathrm{depth}_p(x + \mu, X + \mu) &= \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i + \mu \rangle \le \langle u, x + \mu \rangle\}) \\
&= \min_{\|u\|=1} \#\{i \in [n] : \langle u, x_i \rangle \le \langle u, x \rangle\}) \\
&= \mathrm{depth}_p(x, X)
\end{aligned}$$

end of proof

Fact 2: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

end of proof

Fact 2: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

Proof of Fact 2:

$$\langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle \iff \langle \Sigma^\top u, x_i \rangle \leq \langle \Sigma^\top u, x \rangle$$

end of proof

<u>Fact 2</u>: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

Proof of Fact 2:

$$\begin{aligned} \langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle &\iff \langle \Sigma^\top u, x_i \rangle \leq \langle \Sigma^\top u, x \rangle \\ &\iff \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x_i \rangle \leq \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x \rangle \end{aligned}$$

end of proof

<u>Fact 2</u>: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

Proof of Fact 2:

$$\begin{aligned}
\langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle &\iff \langle \Sigma^\top u, x_i \rangle \leq \langle \Sigma^\top u, x \rangle \\
&\iff \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x_i \rangle \leq \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x \rangle \\
&\iff x_i \in H_{\varphi(u),x} \text{ where } \varphi(u) := \frac{\Sigma^\top u}{\|\Sigma^\top u\|}
\end{aligned}$$

and φ is bijective over $\{u \in \mathbb{R}^p : \|u\| = 1\}$, $\varphi^{-1}(v) = \frac{(\Sigma^{-1})^\top u}{\|(\Sigma^{-1})^\top u\|}$

end of proof

<u>Fact 2</u>: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

Proof of Fact 2:

$$\begin{aligned}
\langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle &\iff \langle \Sigma^\top u, x_i \rangle \leq \langle \Sigma^\top u, x \rangle \\
&\iff \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x_i \rangle \leq \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x \rangle \\
&\iff x_i \in H_{\varphi(u),x} \text{ where } \varphi(u) := \frac{\Sigma^\top u}{\|\Sigma^\top u\|}
\end{aligned}$$

and φ is bijective over $\{u \in \mathbb{R}^p : \|u\| = 1\}$, $\varphi^{-1}(v) = \frac{(\Sigma^{-1})^\top u}{\|(\Sigma^{-1})^\top u\|}$

$$\mathrm{depth}_p(\Sigma x, \Sigma X) = \min_{\|u\|=1} \#\{i \in [n] : \langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle\}$$

end of proof

Fact 2: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

Proof of Fact 2:

$$\begin{aligned} \langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle &\iff \langle \Sigma^\top u, x_i \rangle \leq \langle \Sigma^\top u, x \rangle \\ &\iff \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x_i \rangle \leq \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x \rangle \\ &\iff x_i \in H_{\varphi(u),x} \text{ where } \varphi(u) := \frac{\Sigma^\top u}{\|\Sigma^\top u\|} \end{aligned}$$

and φ is bijective over $\{u \in \mathbb{R}^p : \|u\| = 1\}$, $\varphi^{-1}(v) = \frac{(\Sigma^{-1})^\top u}{\|(\Sigma^{-1})^\top u\|}$

$$\begin{aligned} \mathrm{depth}_p(\Sigma x, \Sigma X) &= \min_{\|u\|=1} \#\{i \in [n] : \langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle\} \\ &= \min_{\|u\|=1} \#\{i \in [n] : x_i \in H_{\varphi(u),x}\} \end{aligned}$$

end of proof

Fact 2: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

Proof of Fact 2:

$$\begin{aligned}
\langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle \iff & \langle \Sigma^\top u, x_i \rangle \leq \langle \Sigma^\top u, x \rangle \\
\iff & \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x_i \rangle \leq \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x \rangle \\
\iff & x_i \in H_{\varphi(u),x} \text{ where } \varphi(u) := \frac{\Sigma^\top u}{\|\Sigma^\top u\|}
\end{aligned}$$

and φ is bijective over $\{u \in \mathbb{R}^p : \|u\| = 1\}$, $\varphi^{-1}(v) = \frac{(\Sigma^{-1})^\top u}{\|(\Sigma^{-1})^\top u\|}$

$$\begin{aligned}
\mathrm{depth}_p(\Sigma x, \Sigma X) &= \min_{\|u\|=1} \#\{i \in [n] : \langle u, \Sigma x_i \rangle \leq \langle u, \Sigma x \rangle\} \\
&= \min_{\|u\|=1} \#\{i \in [n] : x_i \in H_{\varphi(u),x}\} \\
&= \min_{\|u\|=1} \#\{i \in [n] : x_i \in H_{u,x}\}
\end{aligned}$$

end of proof

Fact 2: $\mathrm{depth}_p(\Sigma x, \Sigma X) = \mathrm{depth}_p(x, X)$ for Σ non-singular

Proof of Fact 2:

$$\begin{aligned}
\langle u, \Sigma x_i \rangle \le \langle u, \Sigma x \rangle &\iff \langle \Sigma^\top u, x_i \rangle \le \langle \Sigma^\top u, x \rangle \\
&\iff \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x_i \rangle \le \langle \frac{\Sigma^\top u}{\|\Sigma^\top u\|}, x \rangle \\
&\iff x_i \in H_{\varphi(u),x} \text{ where } \varphi(u) := \frac{\Sigma^\top u}{\|\Sigma^\top u\|}
\end{aligned}$$

and φ is bijective over $\{u \in \mathbb{R}^p : \|u\| = 1\}$, $\varphi^{-1}(v) = \frac{(\Sigma^{-1})^\top u}{\|(\Sigma^{-1})^\top u\|}$

$$\begin{aligned}
\mathrm{depth}_p(\Sigma x, \Sigma X) &= \min_{\|u\|=1} \#\{i \in [n] : \langle u, \Sigma x_i \rangle \le \langle u, \Sigma x \rangle\} \\
&= \min_{\|u\|=1} \#\{i \in [n] : x_i \in H_{\varphi(u),x}\} \\
&= \min_{\|u\|=1} \#\{i \in [n] : x_i \in H_{u,x}\} \\
&= \mathrm{depth}_p(x, X)
\end{aligned}$$

Table of Contents

Maximal depth

Definition

For a dataset X, the **deepest observation level** is

$$k^+(X) = \max_{i=1,\ldots,n} \operatorname{depth}_p(x_i, X)$$

Similarly, the depth of X is the largest depth reached by any point (not necessarily an observation)

$$k^*(X) = \max_{x \in \mathbb{R}^p} \operatorname{depth}_p(x, X)$$

Rem: the points reaching such depth level are affine invariant

Special values

Special values can be reached by the depth of the dataset $k^*(X)$:

- $k^*(X) \leq \frac{n}{2}$
- $k^*(X) = 0$: the dataset X is contained in a (affine) hyperplane : $\{x_1, \ldots, x_n\} \subset \{y \in \mathbb{R}^p : \langle u, y \rangle = \langle u, x \rangle\}$ for some u with $\|u\| = 1$ and $x \in \mathbb{R}^p$
- $k^*(X) = 1$: there is no x_i in the (relative) interior of $\mathrm{conv}(x_1, \ldots, x_n)$, the convex hull of the points $x_1, \ldots, x_n$.

Sub-optimal approach

Definition

$$T_{(k)}(X) = \mathrm{Ave}\{x_i : \mathrm{depth}_p(x_i, X) \geq k\}$$

[3] D. L. Donoho and M. Gasko. “Breakdown properties of location estimates based on halfspace depth and projected outlyingness”. In: *Ann. Statist.* 20.4 (1992), pp. 1803–1827.

Sub-optimal approach

Definition

$$T_{(k)}(X) = \mathrm{Ave}\{x_i : \mathrm{depth}_p(x_i, X) \geq k\}$$

Rem: we write Ave for the averaging operator.

Theorem

The breakdown point for the (affine equivariant) estimator $T_{(k)}(X)$ is bounded by the deepest observation level:

$$\varepsilon^*(T_{(k)}(X), X) = \frac{k^+(X)}{k^+(X) + n}$$

Sketch of proof: put the additional $k^+(X)$ points at the same location, and arbitrary far[3].

[3] D. L. Donoho and M. Gasko. "Breakdown properties of location estimates based on halfspace depth and projected outlyingness". In: *Ann. Statist.* 20.4 (1992), pp. 1803–1827.

Outlyingness measure

Remind the notation: $\langle u, X\rangle = [\langle u, x_1\rangle, \ldots, \langle u, x_n\rangle]$

Definition

The outlyingness of a point x w.r.t. a data set is defined as

$$r_p(x, X) = \max_{\|u\|=1} \frac{|\langle u, x\rangle - \mathrm{Med}_n(\langle u, X\rangle)|}{\mathrm{MAD}_n \langle u, X\rangle}$$

Optimal robust estimator

We say that $x_1, \ldots, x_n \in \mathbb{R}^p$ are in **general position** whenever no more than p points lie in an affine hyperplane (an affine subspace of dimension $p-1$).

Theorem

Provided that $x_1, \ldots, x_n$ are in general position, the estimator

$$\hat{t}_w(X) = \frac{\sum_{i=1}^{n} w(x_i, X) x_i}{\sum_{i=1}^{n} w(x_i, X)}, \quad \text{with weights} \quad w(x_i, X) = \frac{1}{r_p(x_i, X)}$$

is affine equivariant with breakdown point $\varepsilon^* = \frac{n-2p+1}{2n-2p+1}$, if $n \geq 2p$.

Proof: see Donoho (1982)

Computational aspects

The problem of computing $\text{depth}_p(x_1, X)$ NP-hard[4]!

A review for computational challenges computing depth is given in Chen *et al.* (2013), see also Dyckerhoff and Mozharovskyi (2016) and Mozharovskyi(2016)

- 1D : cost is $O(n)$ to compute $\text{depth}_1(x, X)$; simply count how many points are greater/smaller that x in X.
- 2D : cost is[5] $O(n \log(n))$ to compute $\text{depth}_2(x, X)$
- ...

[4] D. S. Johnson and F. P. Preparata. "The densest hemisphere problem". In: *Theoret. Comput. Sci.* 6.1 (1978), pp. 93–107.

[5] P. J. Rousseeuw and I. Ruts. "Algorithm AS 307: Bivariate location depth". In: *J. R. Stat. Soc. Ser. C. Appl. Stat.* 45.4 (1996), pp. 516–526.

Alternative estimator: Minimum Volume Ellipsoid (MVE)[6][7]

Definition

For any constant $c > 0$ and $h \in [n]$, $\hat{t}_n$ is the $\mathrm{MVE}(h, c)$ location estimator (and scatter estimator C_n) are defined by :

$$(\hat{t}_n, \hat{C}_n) \in \underset{t \in \mathbb{R}^p, C \in \mathcal{S}^n_{++}}{\arg\min} \det(C)$$

$$\text{s.t. } \#\{i \in [n] : (x_i - t)^\top C^{-1}(x_i - t) \leq c\} \geq h$$

where $\mathcal{S}^p_{++}$ is the set of positive definite matrices of size p

- the ellipsoid hence created should cover at least h points

[6] P. J. Rousseeuw. "Least median of squares regression". In: *J. Amer. Statist. Assoc.* 79.388 (1984), pp. 871–880.

[7] S. Van Aelst and P. J. Rousseeuw. "Minimum volume ellipsoid". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 1.1 (2009), pp. 71–82.

Properties of MVE

- $\hat{t}_n$ is affine equivariant

[8] H. P. Lopuhaä and P. J. Rousseeuw. "Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices". In: *Ann. Statist.* 19.1 (1991), pp. 229–248.

Properties of MVE

- $\hat{t}_n$ is affine equivariant
- $\hat{C}_n(AX) = A^\top \hat{C}_n A$ for any non singular A

[8] H. P. Lopuhaä and P. J. Rousseeuw. "Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices". In: *Ann. Statist.* 19.1 (1991), pp. 229–248.

Properties of MVE

- $\hat{t}_n$ is affine equivariant
- $\hat{C}_n(AX) = A^\top \hat{C}_n A$ for any non singular A
- $\hat{t}_n$ and $\hat{C}_n$ both have asymptotically $\varepsilon^* = \frac{1}{2}$ when X is in general position[8].

[8] H. P. Lopuhaä and P. J. Rousseeuw. "Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices". In: *Ann. Statist.* 19.1 (1991), pp. 229–248.

References I

- Chen, D., P. Morin, and U. Wagner. "Absolute approximation of Tukey depth: theory and experiments". In: *Comput. Geom.* 46.5 (2013), pp. 566–573.
- Donoho, D. L. "Breakdown properties of multivariate location estimators". PhD thesis. Harvard University, 1982.
- Donoho, D. L. and M. Gasko. "Breakdown properties of location estimates based on halfspace depth and projected outlyingness". In: *Ann. Statist.* 20.4 (1992), pp. 1803–1827.
- Dyckerhoff, R. and P. Mozharovskyi. "Exact computation of the halfspace depth". In: *Computational Statistics & Data Analysis* 98 (2016), pp. 19–30.
- Johnson, D. S. and F. P. Preparata. "The densest hemisphere problem". In: *Theoret. Comput. Sci.* 6.1 (1978), pp. 93–107.
- Lopuhaä, H. P. and P. J. Rousseeuw. "Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices". In: *Ann. Statist.* 19.1 (1991), pp. 229–248.

References II

- Mozharovskyi, P. "Tukey depth: linear programming and applications". In: *arXiv preprint arXiv:1603.00069* (2016).
- Rousseeuw, P. J. "Least median of squares regression". In: *J. Amer. Statist. Assoc.* 79.388 (1984), pp. 871–880.
- Rousseeuw, P. J. and I. Ruts. "Algorithm AS 307: Bivariate location depth". In: *J. R. Stat. Soc. Ser. C. Appl. Stat.* 45.4 (1996), pp. 516–526.
- Tukey, J. W. "Mathematics and the picturing of data". In: *Proceedings of the International Congress of Mathematicians, Vancouver, 1975*. Vol. 2. 1975, pp. 523–531.
- Van Aelst, S. and P. J. Rousseeuw. "Minimum volume ellipsoid". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 1.1 (2009), pp. 71–82.