WEIGHTED DATA DEPTH
AND ITS PROPERTIES

Ondřej Vencálek

Abstract: The paper deals with nonparametric methods for analysis of multivariate data. Generalization of the halfspace depth, so called weighted halfspace depth, is introduced. We consider its properties with respect to the weight function in its definition. One particular property of our interest: points out of the support of probability measure should have depth equal to zero.

1 Introduction

Many statistical problems deal with multivariate random vectors. The classical theory of multivariate data analysis is based on the assumption of multivariate normal distribution of the data. This restriction is often disputable so another, nonparametric, approach should be considered.

One of the most important approaches to the multivariate data is based on so called data depth. The concept of data depth is based on ordering of points in $\mathbb{R}^p$. The principle of ordering is very simple. We can add one real value number to a certain point $x \in \mathbb{R}^p$ that describes its centrality with respect to some probability distribution $P$. This number is called depth of point $x$ with respect to $P$. We denote it $D_P(x)$. The way we add these real numbers to points in $\mathbb{R}^p$ is described by so called depth function. We denote it $D_P : \mathbb{R}^p \rightarrow \mathbb{R}^+$. We order points in $\mathbb{R}^p$ according to their depth.

The J.Tukey’s definition of the halfspace depth function (1975, see [4]) brought quite simple but very natural way of central-outward ordering of points with respect to the probability distribution. The rigorous definition is recalled in Section 2. The approach based on the data depth has been developed since then. A new wave of interest of this approach came in 1990s and has continued so far. Review of this methodology can be found e.g. in [1]. New results are introduced in [2].

Nonparametric statistical inference for multivariate random vectors has been developed based on the data depth. But the definition of the depth function alone is still not unified. Many various depth functions have been introduced.
The most well known are simplicial depth, Oja depth, $L_1$-depth and convex hull peeling depth. General definition of the depth function was proposed by Zuo and Serfling (see [3]) in 2000. This general definition says that the depth function is an arbitrary function which has some desirable properties (we introduce them here without detailed comment: a depth function is any non-negative bounded function which is affine invariant, maximal at the center, monotone relatively to the deepest point and which goes to zero at infinity).

We propose generalization of the halfspace depth. We call it weighted halfspace depth. Its formal definition can be found in Section 2. We investigate its properties. They depend on a weight function that we use for weighted halfspace depth. Not all weight functions are proper. We found some very mild conditions under that the weight function is proper in the sense that all points out of the support of probability have weighted halfspace depth equal to zero. These conditions are presented in Section 3 for a probability distribution with convex support. If we consider nonconvex support of probability measure the conditions are very similar. They are presented in Section 4.

2 Halfspace depth and weighted halfspace depth

In this section we present some basic definitions. Namely the definition of the halfspace depth and the weighted halfspace depth. We also recall the definition of the closed convex support of probability measure.

**Definition 1:** Let $P$ be a probability measure on $\mathbb{R}^p$. The halfspace depth of the point $x \in \mathbb{R}^p$ with respect to $P$ is defined as

$$D_P(x) = \inf_{u : \|u\|=1} P\{y : u^T(y - x) \geq 0\}$$

If we consider empirical probability measure $\hat{P}_n$ instead of $P$ we get a sample version of the halfspace depth. Zuo and Serfling showed (in [3]) that the halfspace depth has all desirable properties of the depth function. One particular property of the halfspace depth is convexity of the areas of the deepest points i.e. for arbitrary $P$ and $d \in (0, 1)$ the set $\{x : D_P(x) \geq d\}$ is convex. This property could be considered as undesirable (e.g. for the uniform distribution on a nonconvex support). Therefore we introduce a generalization of the halfspace depth which has not this property in general. Defining the weighted halfspace depth we follow several consecutive steps. The definition of the halfspace depth can be expressed in the following form:

$$D_P(x) = \inf P\{H : \text{ a closed halfspace: } x \in H\} = \inf \int_H dP,$$

where the infimum is over all closed halfspaces $H$ that contain $x$. Our modification consists in integration of some weight function (w.r.t. $P$). So we
have

\[ p_H = \int_H w(X) dP. \]

Now we are ready to define weighted halfspace depth of a certain point \( x \):

**Definition 2:** Consider \( p \)-dimensional space \( \mathbb{R}^p \), a point \( x = (x_1, \ldots, x_p) \) \( x \in \mathbb{R}^p \) and a probability measure \( P \) on \( \mathbb{R}^p \). Let \( w_+ : \mathbb{R}^p \to [0, \infty) \) be a bounded measurable weight function such that \( w_+(x) = 0 \) for \( x_p < 0 \) and \( w_-(x) = w_-(x_1, \ldots, x_{p-1}, x_p) = w_+(x_1, \ldots, x_{p-1}, -x_p) \).

We define the weighted halfspace depth of the point \( x \) with respect to \( P \) as

\[
D_X(x) = D_P(x) = \inf_{A \in O_p} \mathbb{E}_{P|w_+}(A(X - x)) \mathbb{E}_{P|w_-}(A(X - x)),
\]

(1)

where \( O_p \) denotes the space of all orthogonal \( p \times p \) matrices.

Let’s start with a lemma about two general properties of this depth that hold for any weight function.

**Lemma 1:** The depth function defined by the formula (1) is

- translation invariant, i.e. \( D_{X+a}(x + a) = D_X(x) \)
- rotation invariant, i.e. \( D_{AX}(Ax) = D_X(x) \),

where \( A \) is the matrix of the rotation.

**Proof:** Proof follows directly from the definition. Considering rotation invariancy we just note that \( \{BA : B \in O_p\} = O_p \).

We want to discuss the choice of the weight function in Definition 2. One possible criterion is the property that all points out of the closed convex support of probability measure have depth equal to zero. It means:

\[ x \notin \text{csp}(P) \Rightarrow H_P(x) = 0. \]  

(2)

Recall that the support of a probability measure \( P \) (we denote it \( \text{sp}(P) \)) is the smallest closed set with the probability equal to 1, i.e.

\[ \text{sp}(P) = \bigcap \{ F \in \mathcal{F} : P(F) = 1 \}, \]

where \( \mathcal{F} \) is the class of all closed subsets of \( \mathbb{R}^p \). The closed convex support of probability measure \( P \) is defined as closed convex hull of the support \( \text{sp}(P) \). It is denoted \( \text{csp}(P) \).
3 Convex support of probability measure

Consider a probability distribution with convex support, e.g. the multivariate exponential distribution or the uniform distribution on a convex support. Let’s consider only two-dimensional space ($p = 2$) for simplicity. The following theorem introduces the sufficient condition for property (2):

**Theorem 1:** Let $\text{sp}(P)$ be convex. We denote $W = \{y : w_+(y) > 0\}$. Suppose that

$$\forall x : x_1 = 0, x_2 > 0 \exists U \text{ a neighbourhood of the point } x : U \subset W, \quad (3)$$

holds for the weight function. Then $H_P(x) = 0$ holds for all $x \notin \text{sp}(P)$.

**Proof:** Suppose $x_0 \notin \text{sp}(P)$. We want to prove that its weighted depth will be equal to zero when condition (3) holds.

We denote $x_m = \arg\min \{|x-x_0| : x \in \text{sp}(P)\}$, i.e. $x_m$ is the point of the support of $P$ with the smallest distance from $x_0$. Existence and uniqueness of this point arise from the convexity of $\text{sp}(P)$. We can suppose (without loss of generality) $x_0 = (0,0)$ because of the translation invariance of the weighted depth (see Lemma 1) and we can suppose (without loss of generality) $x_m = (0,v)$, where $v = |x-x_0|$ because of the rotation invariance of the weighted depth.

For such a rotation $\text{sp}(P) \subset H_v \subset H_0$, where $H_v = \{x : x_2 \geq v\}$ and $H_0 = \{x : x_2 \geq 0\}$. We can prove it by an absurdum proof. Suppose that there exist $y \in \text{sp}(P)$ such that $y_2 < v$. Then (from the convexity of $\text{sp}(P)$) all points on the abscissa $x_m$, $y$ are in $\text{sp}(P)$, i.e.

$$x_m + \alpha(y - x_m) \in \text{sp}(P) \quad \forall \alpha \in [0,1].$$

The distance of these points from the origin $(x_0)$ can be expressed as $[(\alpha y_1)^2 + (v + \alpha(y_2 - v))^2]^{1/2}$, what is, for alpha small enough, smaller than $v$. But this is in conflict with the assumption that $x_m = (0,v)$ is the point of $\text{sp}(P)$ with the smallest distance from $x_0 = (0,0)$.

From (3) we have that there exist $U_{x_m}$ a neighbourhood of the point $x_m$ such that $U_{x_m} \subset W$. So we have $U_{x_m} \cap W \cap \text{sp}(P) \neq \emptyset$, hence $E_{Pw_+}(A(X-x_0)) > 0$. It follows from $\text{sp}(P) \subset H_v \subset H_0$ that $E_{Pw_-}(A(X-x_0)) = 0$, so we have $H_P(x_0) = 0$.

□

The class of functions that satisfy the condition (3) from Theorem 1 is still quite broad. Here are some examples of weight functions that satisfies condition (3):

- $w_+(x) > 0$ iff $x_2 \geq 0$, 
Weighted data depth and its properties

- \( w_+(x) > 0 \) iff \( x_2 \geq 0 \) and \( |x_1| \leq h \),
- \( w_+(x) > 0 \) iff \( x_2 \geq 0 \) and \( |x_1| \leq hx_2 \),
- \( w_+(x) > 0 \) iff \( x_2 \geq 0 \) and \( |x_1| \leq h/x_2 \),

where \( h \) is a positive constant. We can see that the condition is really very mild and not too restrictive. However, this condition is not necessary. E.g., consider the uniform distribution on the convex (bounded) support and weight function

\[ w_+(x) > 0 \] if and only if \( x_p \in [0, h] \),

where \( h \) is a positive constant. The condition (3) is not satisfied, but (2) holds.

The form of the condition (3) for general multivariate case \( (p \in \mathbb{N}) \) can be written as follows: For all \( x : x_1 = 0, \ldots, x_{p-1} = 0, x_p > 0 \) there exist \( U \), a neighbourhood of the point \( x \), such that \( U \subseteq W \).

4 Nonconvex support of probability measure

Now we will consider a probability measure \( P \) with nonconvex support. Because of nonconvexity we do not request depth equal to zero for all points out of the support itself, but only for all points out of the closed convex hull of the support \( \text{csp}(P) \). E.g., if we consider the uniform distribution on an annulus, the center of this annulus is the center of symmetry of the distribution, but it is still out of the support. So we do not demand its depth equal to zero.

For a nonconvex support of \( P \) the condition (3) is no more tenable. We can either consider \( \text{sp}(P) \) connected or strengthen the condition on weight function. One possible way how to do it for such a case is introduced in Theorem 2.

**Theorem 2**: Suppose there exists \( n \in \mathbb{N} \) such that \( \text{sp}(P) = \bigcup_{i=1}^{n} K_i \), where \( K_i \) (\( i = 1, \ldots, n \)) is connected subset of \( \mathbb{R}^2 \), and \( \text{sp}(P) \) has no singular point. Denote \( m_{ij} = \min \{|x-y| : x \in K_i, y \in K_j\} \), \( i, j = 1, \ldots, n \).

Consider \( m = \max_{1 \leq i, j \leq n} m_{ij} \).

Let a weight function \( w_+ \) have the following property:

\[ \forall x : |x_1| \leq m/2 \exists U_x \text{ a neighbourhood of } x \text{ such that } U_x \subseteq W. \quad (4) \]

Then \( x \notin \text{csp}(P) \Rightarrow H_P(x) = 0. \)

**Proof**: The proof is very similar to the previous one. We denote the point of \( \text{csp}(P) \) with the smallest distance from \( x_0 = (0, 0) \) by \( x_m = (0, v) \).

We will prove via absurdum proof that there exist a point \( x = (x_1, x_2) \) such that \( |x_1| \leq m/2 \), which is in \( W \) (\( x \in W \)). Suppose that there is
no such a point $x$. Then there must be points $y = (y_1, y_2) \in \text{sp}(P)$ and $z = (z_1, z_2) \in \text{sp}(P)$ such that $y_1 < -m/2$ and $z_1 > m/2$. Hence $|y - z| > m$.

These points are from the different components of connectedness. We take two points from these components with the smallest distance between each other: $y_m$ and $z_m$. For all points $y$ of the one component $y_1 < -m/2$ holds and for all points $z$ of the other component $z_1 > m/2$ holds, we get $|y_m - z_m| > m$, but this is in conflict with the definition of $m$.

From the assumption (4) follows that there exists $U_x$ a neighbourhood of the point $x$ such that $U_x \cap W \cap \text{sp}(P) \neq \emptyset$. Hence (similarly as in proof of the Theorem 1) $H_P(x_0) = 0$.

Note that in a special case $n=1$ (i.e. for nonconvex connected support) we have depth equal to zero for all points out of the closed convex hull of the support when (3) holds.

We have been discussing the depth of the points out of the closed convex hull of the support so far. Now we will discuss properties of the depth of points that are in the convex hull of the support, but out of the support itself. It is easy to show that all points from the closed convex hull of the support have a positive halfspace depth. An advantage of the weighted halfspace depth is that points in the closed convex hull of the support but out of the support itself might have the depth equal to zero.

We explain the advantage on the following example. Consider the uniform distribution on some sector which originates from the circle with the radius $r$ (see Figure 1, part a). All points from the closed convex hull of the support have the halfspace depth greater then zero (Figure 1, part b). Now we consider the weighted halfspace depth with the following weight function

$$w_+(x) = \begin{cases} 1 & \text{if } |x_1| < h, \ x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

where $h$ is some positive constant smaller than $r$. Then the area of points that have the weighted halfspace depth greater then zero is the union of the support and the circle with the same center as the big one, but with radius equal to $h$ (Figure 1, part c). Comparing shapes of the areas with nonzero halfspace depth and nonzero weighted halfspace depth we see that the second one is more similar to the shape of the support of probability measure.
Weighted data depth and its properties

Figure 1: sector-shaped support of the uniform distribution (a); the convex hull of the support i.e. points with nonzero halfspace depth (b); points with nonzero weighted halfspace depth with the weight function (5), where $h = r/2$ (c).

References


Acknowledgement:
The research was supported by the grant GAČR 201/08/0486.

Address: MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

E-mail: vencalek@karlin.mff.cuni.cz