58 COMPUTATION OF ROBUST STATISTICS: DEPTH, MEDIAN, AND RELATED MEASURES

Peter J. Rousseeuw and Mia Hubert

INTRODUCTION

As statistical data sets grow larger and larger, the availability of fast and efficient algorithms becomes ever more important in practice. Classical methods are often easy to compute, even in high dimensions, but they are sensitive to outlying data points. Robust statistics develops methods that are less influenced by abnormal observations, often at the cost of higher computational complexity. Many robust methods, especially those based on ranks, are closely related to geometric or combinatorial problems. An early overview of relations between statistics and geometry was given in [Sha76].

Recently many other (mostly multivariate) statistical methods have been developed that have a combinatorial or geometric character and are computationally intensive. Techniques of computational geometry appear to be very well suited for the development of fast algorithms. Over the last two decades, especially the notion of statistical depth received considerable attention from the computational geometry community. In this chapter we mainly concentrate on depth and multivariate medians, and in Section 58.3 we list other areas of statistics where computational geometry has been of use in constructing efficient algorithms, such as cluster analysis.

58.1 MULTIVARIATE RANKING

A data set consisting of n univariate points is usually ranked in ascending or descending order. Univariate order statistics (i.e., the "kth smallest value out of n") and derived quantities have been studied extensively. The median is defined as the order statistic of rank $(n + 1)/2$ when n is odd, and as the average of the order statistics of ranks $n/2$ and $(n+2)/2$ when n is even. The median and any other order statistic of a univariate data set can be computed in $O(n)$ time. Generalization to higher dimensions is, however, not straightforward.

Alternatively, univariate points may be ranked from the outside inward by assigning the most extreme data points depth 1, the second smallest and second largest data points depth 2, etc. The deepest point then equals the usual median of the sample. The advantage of this type of ranking is that it can be extended to higher dimensions more easily. This section gives an overview of several possible generalizations of depth and the median to multivariate settings. Surveys of statistical applications of multivariate data depth may be found in [LPS99], [ZS00], and [Mos13].

GLOSSARY

Bagdistance: Generalized norm on $\mathbb{R}^d$ based on halfspace depth regions.

Bagplot: Bivariate generalization of the boxplot based on depth regions.

Breakdown value: The smallest fraction of contaminated data points that can move the estimator arbitrarily far away.

Centerpoint: Any point with halfspace depth $\geqslant \lceil n/(d+1) \rceil$.

Deepest fit: Median hyperplane based on regression depth.

Depth: The outside-inward "rank" of a point (not necessarily a data point).

Depth region: The set of all points with depth $\geqslant k$ is called the kth depth region D_k.

Median: The point with maximal depth. When this point is not unique, the median is taken to be the centroid of the depth region with highest depth.

Tukey median: Median based on halfspace depth.

HALFSPACE LOCATION DEPTH

Let $X_n = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be a finite set of data points in $\mathbb{R}^d$. The *Tukey depth* or *halfspace depth* (introduced by [Tuk75] and further developed by [DG92]) of any point $\boldsymbol{\theta}$ in $\mathbb{R}^d$ (not necessarily a data point) determines how central the point is inside the data cloud. The halfspace depth of $\boldsymbol{\theta}$ is defined as the minimal number of data points in any closed halfspace determined by a hyperplane through $\boldsymbol{\theta}$:

$$hdepth(\boldsymbol{\theta}; X_n) = \min_{\|\boldsymbol{u}\|=1} \ \#\{i; \boldsymbol{u}^\tau \boldsymbol{x}_i \geqslant \boldsymbol{u}^\tau \boldsymbol{\theta}\}.$$

Thus, a point lying outside the convex hull of X_n has depth 0, and any data point has depth at least 1. Figure 58.1.1 illustrates this definition for $d = 2$.
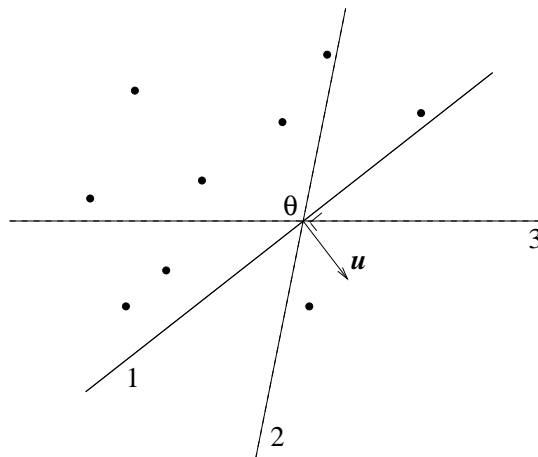

FIGURE 58.1.1

Illustration of the bivariate halfspace depth. Here $\boldsymbol{\theta}$ (which is not a data point itself) has depth 1 because the halfspace determined by $\boldsymbol{u}$ contains only one data point.

The set of all points with depth $\geqslant k$ is called the kth depth region D_k. The halfspace depth regions form a sequence of nested polyhedra. Each D_k is the

intersection of all halfspaces containing at least $n - k + 1$ data points. Moreover, every data point must be a vertex of one or more depth regions. When the innermost depth region is a singleton, that point is called the *Tukey median*. When the innermost depth region is larger than a singleton, the Tukey median is defined as its centroid. This makes the Tukey median unique by construction. Note that the Tukey median does not have to be a data point.

Note that the depth regions give an indication of the shape of the data cloud. Based on this idea one can construct the *bagplot* [RRT99], a bivariate version of the univariate boxplot. Figure 58.1.2 shows such a bagplot. The cross in the white disk is the Tukey median. The dark area is an interpolation between two subsequent depth regions, and contains 50% of the data. This area (the "bag") gives an idea of the shape of the majority of the data cloud. Inflating the bag by a factor of 3 relative to the Tukey median yields the "fence" (not shown), and data points outside the fence are called outliers and marked by stars. Finally, the light gray area is the convex hull of the non-outlying data points.

FIGURE 58.1.2

Bagplot of the heart and spleen size of 73 hamsters.

Bagplot
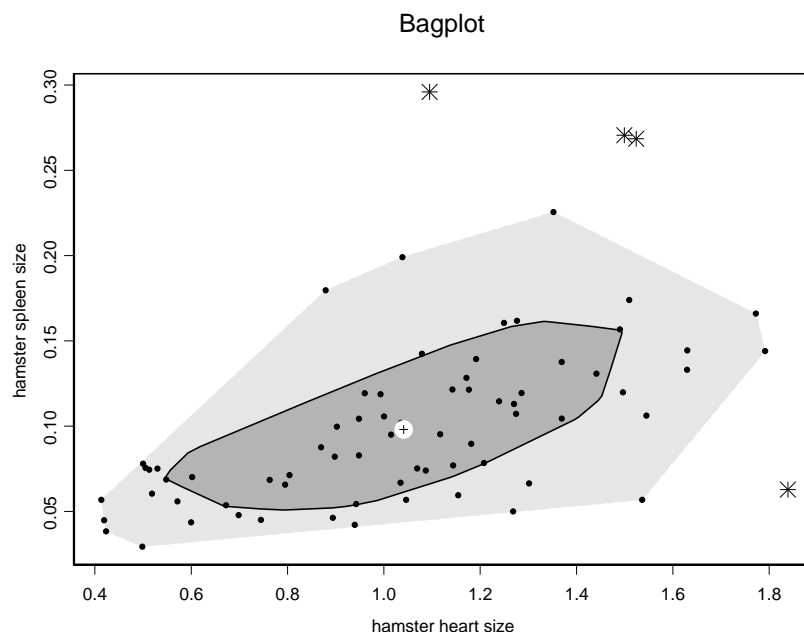
More generally, in the multivariate case one can define the *bagdistance* [HRS15b] of a point x relative to the Tukey median and the bag. Assume that the Tukey median lies in the interior of the bag, not on its boundary (this excludes degenerate cases). Then the bagdistance is the smallest real number λ such that the bag inflated (or deflated) by λ around the Tukey median contains the point x. When the Tukey

median equals $\mathbf{0}$, it is shown in [HRS15b] that the bagdistance satisfies all axioms of a norm except that $||a\boldsymbol{x}|| = |a|||\boldsymbol{x}||$ only needs to hold when $a \geqslant 0$. The bagdistance is used for outlier detection [HRS15a] and statistical classification [HRS15b].

An often used criterion to judge the robustness of an estimator is its *breakdown value.* The breakdown value is the smallest fraction of data points that we need to replace in order to move the estimator of the contaminated data set arbitrarily far away. The classical mean of a data set has breakdown value zero since we can move it anywhere by moving one observation. Note that the breakdown value of any translation equivariant estimator can be at most $1/2$. This can be seen as follows: if we replace half of the data points by a far-away translation image of the remaining half, the estimator cannot know which were the original data.

The Tukey depth and the corresponding median have good statistical properties. The Tukey median T^* is a location estimator with breakdown value $\varepsilon_n(T^*; X_n)$ $\geqslant 1/(d+1)$ for any data set in general position. This means that it remains in a predetermined bounded region unless $n/(d+1)$ or more data points are moved. At an elliptically symmetric distribution the breakdown value becomes $1/3$ for large n, irrespective of d. Moreover, the halfspace depth is invariant under all nonsingular affine transformations of the data, making the Tukey median affine equivariant. Since data transformations such as rotation and rescaling are very common in statistics, this is an important property. The statistical asymptotics of the Tukey median have been studied in [BH99].

The need for fast algorithms for the halfspace depth has only grown over the years, since it is currently being applied to a variety of settings such as nonparametric classification [LCL12]. A related development is the fast growing field of functional data analysis, where the data are functions on a univariate interval (e.g., time or wavelength) or on a rectangle (e.g., surfaces, images). Often the function values are themselves multivariate. One can then define the depth of a curve (surface) by integrating the depth over all points as in [CHSV14]. This functional depth can again be used for outlier detection and classification [HRS15a, HRS15b], but it requires computing depths in many multivariate data sets instead of just one.

CENTERPOINTS

There is a close relationship between the Tukey depth and centerpoints, which have been long studied in computational geometry. In fact, Tukey depth extends the notion of centerpoint. A *centerpoint* is any point with halfspace depth $\geqslant \lceil n/(d+1) \rceil$. A consequence of Helly's theorem is that there always exists at least one centerpoint, so the depth of the Tukey median cannot be less than $\lceil n/(d+1) \rceil$.

OTHER LOCATION DEPTH NOTIONS

1. ***Simplicial depth*** ([Liu90]). The depth of $\boldsymbol{\theta}$ equals the number of simplices formed by $d+1$ data points that contain $\boldsymbol{\theta}$. Formally,

$$sdepth(\boldsymbol{\theta}; X_n) = \#\{(i_1, \ldots, i_{d+1}); \boldsymbol{\theta} \in S[\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_{d+1}}]\}$$

where $S[\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_{d+1}}]$ is the closed simplex with vertices $\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_{d+1}}$. The simplicial median is affine equivariant with a breakdown value bounded above by $1/(d+2)$. Unlike halfspace depth, its depth regions need not be convex, as seen, e.g., in the example in [Mos13].

2. **Oja depth** ([Oja83]). This is also called simplicial volume depth:

$$odepth(\boldsymbol{\theta}; X_n) = \Big(1 + \sum_{(i_1,\ldots,i_d)} \{volume\ S[\boldsymbol{\theta}, \boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_d}]\}\Big)^{-1}.$$

The corresponding median is also affine equivariant, but has zero breakdown value.

3. **Projection depth.** We first define the *outlyingness* ([DG92]) of any point $\boldsymbol{\theta}$ relative to the data set X_n as

$$O(\boldsymbol{\theta}; X_n) = \max_{\|\boldsymbol{u}\|=1} \frac{|\boldsymbol{u}^\tau \boldsymbol{\theta} - \mathrm{med}_i\{\boldsymbol{u}^\tau \boldsymbol{x}_i\}|}{\mathrm{MAD}_i\{\boldsymbol{u}^\tau \boldsymbol{x}_i\}},$$

where the median absolute deviation (MAD) of a univariate data set $\{y_1, \ldots, y_n\}$ is the statistic $\mathrm{MAD}_i\{y_i\} = \mathrm{med}_i|y_i - \mathrm{med}_j\{y_j\}|$. The outlyingness is small for centrally located points and increases if we move toward the boundary of the data cloud. Instead of the median and the MAD, also another pair (T, S) of a location and scatter estimate may be chosen. This leads to different notions of projection depth, all defined as

$$pdepth(\boldsymbol{\theta}; X_n) = (1 + O(\boldsymbol{\theta}; X_n))^{-1}.$$

General projection depth is studied in [Zuo03]. When using the median and the MAD, the projection depth has breakdown value $1/2$ and is affine equivariant. Its depth regions are convex.

4. **Spatial depth** ([Ser02]). Spatial depth is related to multivariate quantiles proposed in [Cha96]:

$$spdepth(\boldsymbol{\theta}; X_n) = 1 - \left\| \frac{1}{n} \sum_{i=1}^n \frac{\boldsymbol{x}_i - \boldsymbol{\theta}}{\|\boldsymbol{x}_i - \boldsymbol{\theta}\|} \right\|$$

The spatial median is also called the L^1 median ([Gow74]). It has breakdown value $1/2$, but is not affine equivariant (it is only equivariant with respect to translations, multiplication by a scalar factor, and orthogonal transformations). For a recent survey on the computation of the spatial median see [FFC12].

A comparison of the main properties of the different location depth medians is given in Table 58.1.1.

ARRANGEMENT AND REGRESSION DEPTH

Following [RH99b] we now define the depth of a point relative to an arrangement of hyperplanes (see Chapter 28). A point $\boldsymbol{\theta}$ is said to have zero **arrangement depth** if there exists a ray $\{\boldsymbol{\theta} + \lambda \boldsymbol{u}; \lambda \geqslant 0\}$ that does not cross any of the hyperplanes h_i in the arrangement. (A hyperplane parallel to the ray is counted as intersecting at infinity.) The arrangement depth of any point $\boldsymbol{\theta}$ is defined as the minimum number of hyperplanes intersected by any ray from $\boldsymbol{\theta}$. Figure 58.1.3 shows an arrangement of lines. In this plot, the points $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ have arrangement depth 0 and the point

TABLE 58.1.1   Comparison of several location depth medians.

MEDIAN	BREAKDOWN VALUE	AFFINE EQUIVARIANCE
Tukey	worst-case $1/(d+1)$ typically $1/3$	yes
Simplicial	$\leqslant 1/(d+2)$	yes
Oja	$2/n \approx 0$	yes
Projection	$1/2$	yes
Spatial	$1/2$	no

$\boldsymbol{\xi}$ has arrangement depth 2. The arrangement depth is always constant on open cells and on cell edges. It was shown ([RH99b]) that any arrangement of lines in the plane encloses a point with arrangement depth at least $\lceil n/3 \rceil$, giving rise to a new type of "centerpoints."

FIGURE 58.1.3

Example of arrangement depth. In this arrangement of lines, the points $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ have arrangement depth 0, whereas $\boldsymbol{\xi}$ has arrangement depth 2. (See Figure 58.1.4 for the dual plot.)
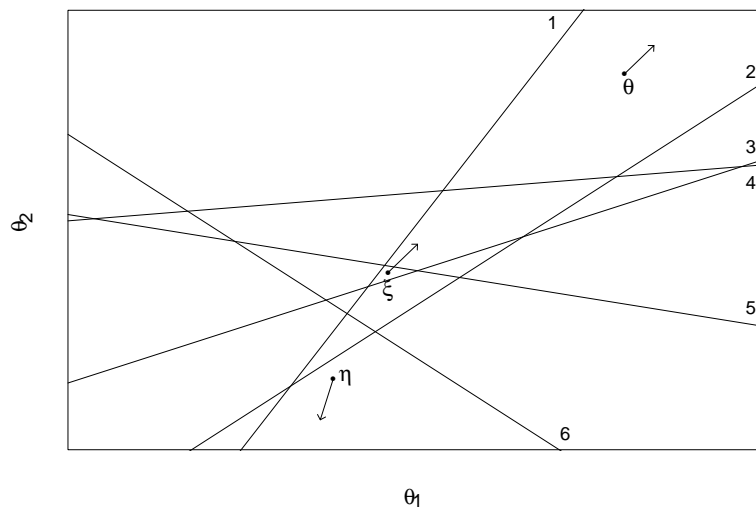

This notion of depth was originally defined ([RH99]) in the dual, as the depth of a regression hyperplane $H_{\boldsymbol{\theta}}$ relative to a point configuration of the form $Z_n = \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_n, y_n)\}$ in $\mathbb{R}^{d+1}$. **Regression depth** ranks hyperplanes according to how well they fit the data in a regression model, with $\boldsymbol{x}$ containing the predictor variables and y the response. A vertical hyperplane (given by $\boldsymbol{a}^{\tau}\boldsymbol{x} = \text{constant}$), which cannot be used to predict future response values, is called a "nonfit" and assigned regression depth 0. The regression depth of a general hyperplane $H_{\boldsymbol{\theta}}$ is

found by rotating H_θ in a continuous movement until it becomes vertical. The minimum number of data points that is passed in such a rotation is called the regression depth of H_θ. Figure 58.1.4 is the dual representation of Figure 58.1.3. (For instance, the line $\boldsymbol{\theta}$ has slope θ_1 and intercept θ_2 and corresponds to the point (θ_1, θ_2) in Figure 58.1.3.) The lines $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ have regression depth 0, whereas the line $\boldsymbol{\xi}$ has regression depth 2.

FIGURE 58.1.4

Example of the regression depth of a line in a bivariate configuration of points. The lines $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ have regression depth 0, whereas the line $\boldsymbol{\xi}$ has regression depth 2. (This is the dual of Figure 58.1.3.)
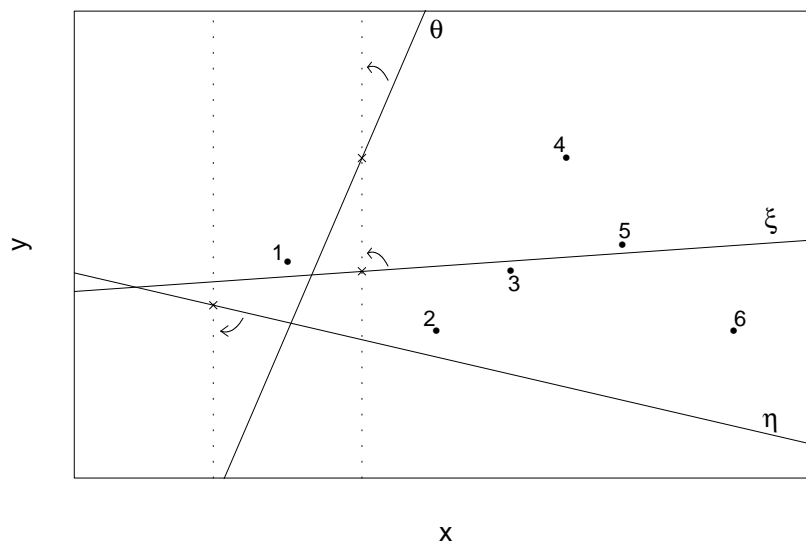
In statistics one is interested in the *deepest fit* or regression depth median, because this is a line (hyperplane) about which the data are well-balanced. The statistical properties of regression depth and the deepest fit are very similar to those of the Tukey depth and median. The bounds on the maximal depth are almost the same. Moreover, for both depth notions the value of the maximal depth can be used to characterize the symmetry of the distribution ([RS04]). The breakdown value of the deepest fit is at least $1/(d+1)$ and under linearity of the conditional median of y given $\boldsymbol{x}$ it converges to $1/3$. In the next section, we will see that the optimal complexities for computing the depth and the median are also comparable to those for halfspace depth. For a detailed comparison of the properties of halfspace and regression depth, see [HRV01].

The arrangement depth region D_k is defined in the primal, as the set of points with arrangement depth at least k. Contrary to the Tukey depth, these depth regions need not be convex. But nevertheless it was proved that there always exists a point with arrangement depth at least $\lceil n/(d+1) \rceil$ ([ABE$^+$00]). An analysis-based proof was given in [Miz02].

ARRANGEMENT LEVELS

Arrangement depth is undirected (isotropic) in the sense that it is defined as a minimum over all possible directions. If we restrict ourselves to vertical directions $\boldsymbol{u}$ (i.e., up or down), we obtain the usual levels of the arrangement (cf. Section 28.2). The absence of preferential directions makes arrangement depth invariant under affine transformations.

58.2 COMPUTING DEPTH

Although the definitions of depth are intuitive, the computational aspects can be quite challenging. Below we will first focus on the bivariate case, and then on higher dimensions.

Algorithms for depth-related measures are often more complex for data sets which are not in general position than for data sets which are. For example, the boundaries of subsequent halfspace depth regions are always disjoint when the data are in general position, but this does not hold for nongeneral position. Preferably, algorithms should be able to handle both the general position case and the nongeneral position case directly. As a quick fix, algorithms which were made for general position can also be applied in the other case if one first adds small random errors to the data points. This is a standard perturbation technique sometimes referred to as dithering. For large data sets, dithering has only a limited effect on the results.

BIVARIATE ALGORITHMS

For the bivariate case several algorithms have been developed early on. Table 58.2.1 gives an overview of algorithms, each of which has been implemented, to compute the depth in a given point $\boldsymbol{\theta}$ in $\mathbb{R}^2$. These algorithms are time-optimal, since the problem of computing these bivariate depths has an $\Omega(n \log n)$ lower bound ([ACG$^+$02], [LS03a]).

The algorithms for halfspace and simplicial depth are based on the same technique. First, data points are radially sorted around $\boldsymbol{\theta}$. Then a line through $\boldsymbol{\theta}$ is rotated. The depth is calculated by counting the number of points that are passed by the rotating line in a specific manner. The planar arrangement depth algorithm is easiest to visualize in the regression setting. To compute the depth of a hyperplane $H_{\boldsymbol{\theta}}$ with coefficients $\boldsymbol{\theta}$, the data are first sorted along the x-axis. A vertical line L is then moved from left to right and each time a data point is passed, the number of points above and below $H_{\boldsymbol{\theta}}$ on both sides of L is updated.

In general, computing a median is harder than computing the depth in a point, because typically there are many candidate points. For instance, for the bivariate simplicial median the currently best algorithm requires $O(n^4)$ time, whereas its corresponding depth needs only $O(n \log n)$. The simplicial median seems difficult to compute because there are $O(n^4)$ candidate points (namely, all intersections of lines passing through two data points) and the simplicial depth regions have irregular shapes, but of course a faster algorithm may yet be found.

Fortunately, in several important cases the median can be computed without computing the depth of individual points. A linear-time algorithm to compute a

TABLE 58.2.1   Computing the depth of a bivariate point.

DEPTH	TIME COMPLEXITY	SOURCE
Tukey depth	$O(n \log n)$	[RR96a]
Simplicial depth	$O(n \log n)$	[RR96a]
Arrangement/regression depth	$O(n \log n)$	[RH99]

bivariate centerpoint was described in [JM94]. Table 58.2.2 gives an overview of algorithms to compute bivariate depth-based medians. For the bivariate Tukey median the lower bound $\Omega(n \log n)$ was proved in [LS00], and the currently fastest algorithm takes $O(n \log^3 n)$ time ([LS03a]). The lower bound $\Omega(n \log n)$ also holds for the median of arrangement (regression) depth as shown by [LS03b]. Fast algorithms were devised by [LS03b] and [KMR$^+$08].

TABLE 58.2.2   Computing the bivariate median.

MEDIAN	TIME COMPLEXITY	SOURCE
Tukey median	$O(n \log^3 n)$	[LS03a]
Simplicial median	$O(n^4)$	[ALS$^+$03]
Oja median	$O(n \log^3 n)$	[ALS$^+$03]
Regression depth median	$O(n \log n)$	[LS03b]

The computation of bivariate halfspace depth regions has also been studied. The first algorithm [RR96b] required $O(n^2 \log n)$ time per depth region. An algorithm to compute all regions in $O(n^2)$ time is constructed and implemented in [MRR$^+$03]. This algorithm thus also yields the Tukey median. It is based on the dual arrangement of lines where topological sweep is applied. A completely different approach is implemented in [KMV02]. They make direct use of the graphics hardware to approximate the depth regions of a set of points in $O(nW + W^3) + nCW^2/512$ time, where the pixel grid is of dimension $(2W + 1) \times (2W + 1)$. Recently, [BRS11] constructed an algorithm to update halfspace depth and its regions when points are added to the data set.

ALGORITHMS IN HIGHER DIMENSIONS

The calculation of depth regions and medians is more computationally intensive in higher dimensions. In statistical practice such data are quite common, and therefore reliable and efficient algorithms are needed. The computational aspects of depth in higher dimensions are still being explored.

The first algorithms to compute the halfspace and regression depth of a given point in $\mathbb{R}^d$ with $d > 2$ were constructed in [RS98] and require $O(n^{d-1} \log n)$ time. The main idea was to use projections onto a lower-dimensional space. This reduces the problem to computing bivariate depths, for which the existing algorithms have optimal time complexity. In [BCI$^+$08] theoretical output-sensitive algorithms for

the halfspace depth are proposed. An interesting computational connection between halfspace depth and multivariate quantiles was provided in [HPS10] and [KM12]. More recently, [DM16] provided a generalized version of the algorithm of [RS98] together with C++ code. For the depth regions of halfspace depth in higher dimensions an algorithm was recently proposed in [LMM14].

For the computation of projection depth see [LZ14]. The simplicial depth of a point in $\mathbb{R}^3$ can be computed in $O(n^2)$ time, and in $\mathbb{R}^4$ the fastest algorithm needs $O(n^4)$ time [CO01]. For higher dimensions, no better algorithm is known than the straightforward $O(n^{d+1})$ method to compute all simplices.

When the number of data points and dimensions are such that the above algorithms become infeasible, one can resort to approximate algorithms. For halfspace depth such approximate algorithms were proposed in [RS98] and [CMW13]. An approximate algorithm in $\mathbb{R}^3$ based on a randomized data structure was proposed in [AC09]. An approximation to the Tukey median using steepest descent can be found in [SR00]. In [VRHS02] an algorithm is described to approximate the deepest regression fit in any dimension. A randomized algorithm for the Tukey median was proposed in [Cha04].

58.3 OTHER STATISTICAL TECHNIQUES

Computational geometry has provided fast and reliable algorithms for many other statistical techniques.

Linear regression is a frequently used statistical technique. The ordinary least squares regression, minimizing the sum of squares of the residuals, is easy to calculate, but produces unreliable results whenever one or more outliers are present in the data. Robust alternatives are often computationally intensive. We here give some examples of regression methods for which geometric or combinatorial algorithms have been constructed.

1. L^1 *regression.* This well-known alternative to least squares regression minimizes the sum of the absolute values of the residuals, and is robust to vertical outliers. Algorithms for L^1 regression may be found in, e.g., [YKII88] and [PK97].

2. **Least median of squares (LMS) regression** ([Rou84]). This method minimizes the median of the squared residuals and has a breakdown value of $1/2$. To compute the bivariate LMS line, an $O(n^2)$ algorithm using topological sweep has been developed [ES90]. An approximation algorithm for the LMS line was constructed in [MNR$^+$97]. The recent algorithm of [BM14] uses mixed integer optimization.

3. **Median slope regression** ([The50], [Sen68]). This bivariate regression technique estimates the slope as the median of the slopes of all lines through two data points. An algorithm with optimal complexity $O(n \log n)$ is given in [BC98], and a more practical randomized algorithm in [DMN92].

4. **Repeated median regression** ([Sie82]). Median slope regression takes the median over all couples (d-tuples in general) of data points. Here, this median is replaced by d nested medians. For the bivariate repeated median regression line, [MMN98] provide an efficient randomized algorithm.

The aim of cluster analysis (Sections 48.5 and 54.1) is to divide a data set into clusters of similar objects. Partitioning methods divide the data into k groups. Hierarchical methods construct a tree (called a dendrogram) such that each cut of the tree gives a partition of the data set. A selection of clustering methods with accompanying algorithms is presented in [SHR97]. The general problem of partitioning a data set into groups such that the partition minimizes a given error function f is NP-hard. However, for some special cases efficient algorithms exist. For a small number of clusters in low dimensions, exact algorithms for partitioning methods can be constructed. Constructing clustering trees is also closely related to geometric problems (see e.g., [Epp97], [Epp98]).

58.4 SOURCES AND RELATED MATERIAL

SURVEYS

All results not given an explicit reference above may be traced in these surveys.

[Mos13]: A survey of multivariate data depth and its statistical applications.

[Sha76]: An overview of the computational complexities of basic statistics problems like ranking, regression, and classification.

[Sma90]: An overview of several multivariate medians and their basic properties.

[ZS00]: A classification of multivariate data depths based on their statistical properties.

RELATED CHAPTERS

Chapter 1: Finite point configurations
Chapter 28: Arrangements
Chapter 54: Pattern recognition

REFERENCES

[ABE+00] N. Amenta, M. Bern, D. Eppstein, and S.-H. Teng. Regression depth and center points. *Discrete Comput. Geom.*, 23:305–323, 2000.

[AC09] P. Afshani and T.M. Chan. On approximate range counting and depth. *Discrete Comput. Geom.*, 42:3–21, 2009.

[ACG+02] G. Aloupis, C. Cortés, F. Gómez, M. Soss, and G.T. Toussaint. Lower bounds for computing statistical depth. *Comput. Stat. Data Anal.*, 40:223–229, 2002.

[ALS+03] G. Aloupis, S. Langerman, M. Soss, and G.T. Toussaint. Algorithms for bivariate medians and a Fermat-Torricelli problem for lines. *Comput. Geom.*, 26:69–79, 2003.

[BC98] H. Brönnimann and B. Chazelle. Optimal slope selection via cuttings. *Comput. Geom.*, 10:23–29, 1998.

[BCI+08]   D. Bremner, D. Chen, J. Iacono, S. Langerman, and P. Morin. Output-sensitive algorithms for Tukey depth and related problems. *Stat. Comput.*, 18:259–266, 2008.

[BH99]   Z.-D. Bai and X. He. Asymptotic distributions of the maximal depth estimators for regression and multivariate location. *Ann. Statist.*, 27:1616–1637, 1999.

[BM14]   D. Bertsimas and R. Mazumder. Least quantile regression via modern optimization. *Ann. Statist.*, 42:2494–2525, 2014.

[BRS11]   M. Burr, E. Rafalin, and D.L. Souvaine. Dynamic maintenance of halfspace depth for points and contours. Preprint, `arXiv:1109.1517`, 2011.

[Cha96]   P. Chaudhuri. On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.*, 91:862–872, 1996.

[Cha04]   T.M. Chan. An optimal randomized algorithm for maximum Tukey depth. In *Proc. ACM-SIAM Sympos. Discrete Algorithms*, 430–436, 2004.

[CHSV14]   G. Claeskens, M. Hubert, L. Slaets, and K. Vakili. Multivariate functional halfspace depth. *J. Amer. Statist. Assoc.*, 109:411–423, 2014.

[CMW13]   D. Chen, P. Morin, and U. Wagner. Absolute approximation of Tukey depth: Theory and experiments. *Comput. Geom.*, 46:566–573, 2013.

[CO01]   A.Y. Cheng and M. Ouyang. On algorithms for simplicial depth. In *Proc. 13th Canadian Conf. on Comp. Geom.*, pages 53–56, Waterloo, 2001.

[DG92]   D.L. Donoho and M. Gasko. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Ann. Statist.*, 20:1803–1827, 1992.

[DM16]   R. Dyckerhoff and P. Mozharovskyi. Exact computation of the halfspace depth. *Comput. Statist. Data Anal.*, 98:19–30, 2016.

[DMN92]   M.B. Dillencourt, D.M. Mount, and N.S. Netanyahu. A randomized algorithm for slope selection. *Internat. J. Comput. Geom. Appl.*, 2:1–27, 1992.

[Epp97]   D. Eppstein. Faster construction of planar two-centers. In *Proc. 8th Annu. ACM-SIAM Sympos. Discrete Algorithms*, pages 131–138, 1997.

[Epp98]   D. Eppstein. Fast hierarchical clustering and other applications of dynamic closest pairs. In *Proc. 9th ACM-SIAM Sympos. Discrete Algorithms*, pages 619–628, 1998.

[ES90]   H. Edelsbrunner and D.L. Souvaine. Computing least median of squares regression lines and guided topological sweep. *J. Amer. Statist. Assoc.*, 85:115–119, 1990.

[FFC12]   H. Fritz, P. Filzmoser, and C. Croux. A comparison of algorithms for the multivariate L_1-median. *Comput. Statist.*, 27:393–410, 2012.

[Gow74]   J.C. Gower. The mediancenter. *J. Roy. Statist. Soc. Ser. C*, 32:466–470, 1974.

[HPS10]   M. Hallin, D. Paindaveine, and M. Šiman. Multivariate quantiles and multiple-output regression quantiles: From L_1-optimization to halfspace depth. *Ann. Statist.*, 38:635–669, 2010.

[HRV01]   M. Hubert, P.J. Rousseeuw, and S. Van Aelst. Similarities between location depth and regression depth. In L.T. Fernholz, editor, *Statistics in Genetics and in the Environmental Sciences*, pages 153–162, Birkhäuser, Basel, 2001.

[HRS15a]   M. Hubert, P.J. Rousseeuw, and P. Segaert. Multivariate functional outlier detection. *Stat. Methods Appl.*, 24:177–202, 2015.

[HRS15b]   M. Hubert, P.J. Rousseeuw, and P. Segaert. Multivariate and functional classification using depth and distance. *Adv. Data Anal. Classif.*, online first, 2016.

[JM94]   S. Jadhav and A. Mukhopadhyay. Computing a centerpoint of a finite planar set of points in linear time. *Discrete Comput. Geom.*, 12:291–312, 1994.

[KM12]   L. Kong and I. Mizera. Quantile tomography: Using quantiles with multivariate data. *Statist. Sinica*, 22:1589–1610, 2012.

[KMR$^+$08] M. van Kreveld, J.S.B. Mitchell, P. Rousseeuw, M. Sharir, J. Snoeyink, and B. Speckmann. Efficient algorithms for maximum regression depth. *Discrete Comput. Geom.*, 39:656–677, 2008.

[KMV02]   S. Krishnan, N.H. Mustafa, and S. Venkatasubramanian. Hardware-assisted computation of depth contours. In *Proc. 13th ACM-SIAM Sympos. Discrete Algorithms*, pages 558–567, 2002.

[LS00]   S. Langerman and W. Steiger. Computing a maximal depth point in the plane. In *Proc. Japan Conf. Discrete Comput. Geom.*, pages 46–47, 2000.

[LS03a]   S. Langerman and W. Steiger. Optimization in arrangements. In *Proc. 20th Sympos. Theor. Aspects Comp. Sci.*, vol. 2607 of *LNCS*, pages 50–61, Springer, Berlin, 2003.

[LS03b]   S. Langerman and W. Steiger. The complexity of hyperplane depth in the plane. *Discrete Comput. Geom.*, 30:299–309, 2003.

[LCL12]   J. Li, J.A. Cuesta-Albertos, and R.Y. Liu. DD-classifier: Nonparametric classification procedure based on DD-plot. *J. Amer. Statist. Assoc.*, 107:737–753, 2012.

[Liu90]   R.Y. Liu. On a notion of data depth based on random simplices. *Ann. Statist.*, 18:405–414, 1990.

[LMM14]   X. Liu, K. Mosler, and P. Mozharovskyi. Fast computation of Tukey trimmed regions in dimension $p > 2$. arXiv:1412.5122 [stat.CO], 2014.

[LPS99]   R.Y. Liu, J. Parelius, and K. Singh. Multivariate analysis by data depth: descriptive statistics, graphics and inference. *Ann. Statist.*, 27:783–840, 1999.

[LZ14]   X. Liu and Y. Zuo. Computing projection depth and its associated estimators. *Stat. Comput.*, 24:51–63, 2014.

[Miz02]   I. Mizera. On depth and deep points: a calculus. *Ann. Statist.* 30:1681–1736, 2002.

[MMN98]   J. Matoušek, D.M. Mount, and N.S. Netanyahu. Efficient randomized algorithms for the repeated median line estimator. *Algorithmica*, 20:136–150, 1998.

[MNR$^+$97] D.M. Mount, N.S. Netanyahu, K. Romanik, R. Silverman, and A.Y. Wu. A practical approximation algorithm for the LMS line estimator. In *Proc. 8th Annu. ACM-SIAM Sympos. Discrete Algorithms*, pages 473–482, 1997.

[Mos13]   K. Mosler. Depth statistics. In C. Becker, R. Fried, and S. Kuhnt, editors, *Robustness and Complex Data Structures*, pages 17–34, Springer, Heidelberg, 2013.

[MRR$^+$03] K. Miller, S. Ramaswami, P. Rousseeuw, J.A. Sellarès, D.L. Souvaine, I. Streinu, and A. Struyf. Efficient computation of location depth contours by methods of computational geometry. *Stat. Comput.*, 13:153–162, 2003.

[Oja83]   H. Oja. Descriptive statistics for multivariate distributions. *Statist. Probab. Lett.*, 1:327–332, 1983.

[PK97]   S. Portnoy and R. Koenker. The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators. *Stat. Sci.*, 12:279–300, 1997.

[RH99]   P.J. Rousseeuw and M. Hubert. Regression depth. *J. Amer. Statist. Assoc.*, 94:388–402, 1999.

[RH99b]   P.J. Rousseeuw and M. Hubert. Depth in an arrangement of hyperplanes. *Discrete Comput. Geom.*, 22:167–176, 1999.

[Rou84]   P.J. Rousseeuw. Least median of squares regression. *J. Amer. Statist. Assoc.*, 79:871–880, 1984.

[RR96a]   P.J. Rousseeuw and I. Ruts. Algorithm AS 307: Bivariate location depth. *J. Roy. Statist. Soc. Ser. C*, 45:516–526, 1996.

[RR96b]   I. Ruts and P.J. Rousseeuw. Computing depth contours of bivariate point clouds. *Comput. Stat. Data Anal.*, 23:153–168, 1996.

[RRT99]   P.J. Rousseeuw, I. Ruts, and J.W. Tukey. The bagplot: A bivariate boxplot. *Amer. Statist.*, 53:382–387, 1999.

[RS98]   P.J. Rousseeuw and A. Struyf. Computing location depth and regression depth in higher dimensions. *Stat. Comput.*, 8:193–203, 1998.

[RS04]   P.J. Rousseeuw and A. Struyf. Characterizing angular symmetry and regression symmetry. *J. Statist. Plann. Inference*, 122:161-173, 2004.

[Sen68]   P.K. Sen. Estimates of the regression coefficient based on Kendall's tau. *J. Amer. Statist. Assoc.*, 63:1379–1389, 1968.

[Ser02]   R. Serfling. A depth function and a scale curve based on spatial quantiles. In Y. Dodge, editor, *Statistical Data Analysis Based on the L_1-Norm and Related Methods*, pages 25–38, Birkhaüser, Basel, 2002.

[Sha76]   M.I. Shamos. Geometry and statistics: Problems at the interface. In J.F. Traub, editor, *Algorithms and Complexity: New Directions and Recent Results*, pages 251–280, Academic Press, Boston, 1976.

[SHR97]   A. Struyf, M. Hubert, and P.J. Rousseeuw. Integrating robust clustering techniques in S-PLUS. *Comput. Stat. Data Anal.*, 26:17–37, 1997.

[Sie82]   A. Siegel. Robust regression using repeated medians. *Biometrika*, 69:242–244, 1982.

[Sma90]   C.G. Small. A survey of multidimensional medians. *Internat. Statistical Review*, 58:263–277, 1990.

[SR00]   A. Struyf and P.J. Rousseeuw. High-dimensional computation of the deepest location. *Comput. Stat. Data Anal.*, 34:415–426, 2000.

[The50]   H. Theil. A rank-invariant method of linear and polynomial regression analysis (parts 1-3). *Nederl. Akad. Wetensch. Ser. A*, 53:386–392, 521–525, 1397–1412, 1950.

[Tuk75]   J.W. Tukey. Mathematics and the picturing of data. In *Proc. Internat. Congr. of Math.*, 2, pages 523–531, Vancouver, 1975.

[VRHS02]   S. Van Aelst, P.J. Rousseeuw, M. Hubert, and A. Struyf. The deepest regression method. *J. Multivar. Anal.*, 81:138–166, 2002.

[YKII88]   P. Yamamoto, K. Kato, K. Imai, and H. Imai. Algorithms for vertical and orthogonal L^1 linear approximation of points. In *Proc. 4th Sympos. Comput. Geom.*, pages 352–361, ACM Press, 1988.

[Zuo03]   Y. Zuo. Projection based depth functions and associated medians. *Ann. Statist.*, 31:1460–1490, 2003.

[ZS00]   Y. Zuo and R. Serfling. General notions of statistical depth function. *Ann. Statist.*, 28:461–482, 2000.