Calculation of simplicial depth estimators for polynomial regression with application to tests in shape analysis

R. Wellmann [a,*], S. Katina [b,1], Christine H. Müller [a]

[a]*Department of Mathematics, Carl von Ossietzky University, Postfach 2503, D-26111 Oldenburg, Germany*

[b]*Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and CS, Comenius University, Mlynská dolina, 842 48 Bratislava, Slovakia*

Abstract

A fast algorithm for calculating the simplicial depth of a single parameter value of a polynomial regression model and for calculating the maximum simplicial depth within an affine subspace of the parameter space or a polyhedron is presented. Since the maximum simplicial depth estimator is not unique, l_1 and l_2 methods are used to make the estimator unique. This estimator is compared with other estimators in examples of linear and quadratic regression. Furthermore, it is shown how the maximum simplicial depth can be used to derive distribution-free asymptotic α-level tests for testing hypotheses in polynomial regression models. The tests are applied on a problem of shape analysis where it is tested how the relative head length of the fish species *Lepomis gibbosus* depends on the size of these fishes. It is also tested whether the dependency can be described by the same polynomial regression function within different populations.

Key words: polynomial regression, simplicial depth, maximum depth estimator, distribution free tests, one-sample tests, two-sample tests, shape analysis
PACS: AMS Subject classification: 62 G 05, 62 G 10, 62 J 05, 62 J 12, 62 G 20

* Corresponding author
Email addresses: `robin.wellmann@mail.uni-oldenburg.de` (R. Wellmann), `katina@fmph.uniba.sk` (S. Katina), `mueller@math.uni-oldenburg.de` (Christine H. Müller).

1 Introduction

Many distribution-free tests as the Wilcoxon's rank sum test base on ranks. Ranks are defined for univariate location data. One possibility to extend the notion of ranks to multivariate data and regression is the concept of data depth. Maximizing the data depth leads also to new outlier robust estimators.

Different depth notions for multivariate data are, for example, the half space depth of Tukey (1975) and the simplicial depth of Liu (1988, 1990). For other depth notions see the book of Mosler (2002) and the references in it. Multivariate depth concepts were transferred to regression by Rousseeuw and Hubert (1999) and studied in a more general context by Mizera (2002).

Although data depth generalizes the notion of rank, data depth is mainly used for estimation. Only few papers deal with tests based on data depth. Liu (1992) and Liu and Singh (1993) proposed rather special distribution-free multivariate rank tests based on depth notions. Distribution-free tests for regression based on depth notions were derived by Van Aelst et al. (2002), Müller (2005) and Wellmann et al. (2006). Van Aelst et al. (2002) even derived an exact test based on the regression depth of Rousseeuw and Hubert (1999) but did it only for linear regression. Regarding simplicial regression depth, Müller (2005) derived the asymptotic distribution of the test statistic not only for linear regression but also for quadratic regression. This approach was extended to polynomial regression of arbitrary degree and multiple regression through the origin by Wellmann et al. (2006). Simplicial regression depth is related to the regression depth of Rousseeuw and Hubert (1999) as the simplicial depth of Liu (1988, 1990) to the half space depth of of Tukey (1975) for multivariate location. The advantage of a simplicial depth notion is that the depth function is an U-statistic and that therefore the asymptotic distribution is in principle known.

However, the calculation of simplicial regression depth is not that easy. In this paper we propose an algorithm to calculate the simplicial regression depth of a single parameter value of a polynomial regression model of arbitrary degree. Moreover we present an algorithm for calculating the maximum simplicial regression depth within an affine subspace of the parameter space or within a polyhedron of the parameter space. Such subsets of the parameter space are of interest in hypotheses testing so that these algorithms are in particular useful for testing hypotheses in polynomial regression models. But the algorithms are also useful for calculating outlier robust estimators based on the maximum simplicial regression depth. Since the maximum simplicial depth estimator, which is the parameter which maximizes the simplicial depth, is not unique in the general case, we propose two methods based on l_1 and l_2 minimization to obtain unique estimators.

In Section 2, the theoretical background of the simplicial regression depth is provided. To improve the tests, the regression depth of Rousseeuw and Hubert (1999) which was used in Müller (2005) to define the simplicial regression depth is modified to a harmonized depth. Based on this harmonized regression depth, the simplicial regression depth and the maximum simplicial depth estimator for polynomial regression is defined in Section 2. Section 3 provides the algorithms for calculating the simplicial depth and the maximum simplicial depth estimator for polynomial regression. It also shows how to find a unique maximum simplicial depth estimator. In Section 4 the distribution-free tests based on simplicial depth for one-sample problems given by Wellmann et al. (2006) are provided and extended to general two-sample tests for polynomial regression. The Sections 5 and 6 provide applications. In Section 5 the maximum simplicial depth estimator is compared with other estimators by means of the Hertzsprung-Russell data. The new distribution free tests are applied on some hypotheses appearing in a problem of shape analysis of the fish species *Lepomis gibbosus* in Section 6. There it is also shown that hypotheses which are difficult to test in classical theory can be tested with the new tests.

2   The theoretical background

2.1   The simplicial depth

We assume that the bivariate random variables $Z_1, ..., Z_N$ are independent and identically distributed throughout the paper. The variables Z_n have values in $\mathcal{Z} \subset \mathbb{R}^2$. We assume, that there is a known family of probability measures $\mathcal{P} = \{P_\theta^{(Z_1,...,Z_N)} : \theta \in \Theta\}$ with $\Theta = \mathbb{R}^q$. For given observations $z_1, ..., z_N \in \mathcal{Z}$, we always write $z = (z_1, ..., z_N)$ and $z_n = (y_n, t_n)$.

Mizera (2002) defined depth via general loss functions while Müller (2005) regarded likelihood depth where the loss is given by the negative likelihood function. In most models for polynomial regression, the likelihood functions have special properties. We call functions with these properties "quality functions for polynomial regression".

3

Definition 2.1 (Quality functions for polynomial regression) *Let* x : $\mathbb{R} \to \mathbb{R}^q$ *be the function with* $x(t) = (1, t, ..., t^{q-1})^T$ *and for* $z_0 = (y_0, t_0) \in \mathcal{Z}$ *take* $\varphi_{z_0} : \mathbb{R} \to \mathbb{R}$ *to be a function with continuous derivatives and exactly one critical point* $h(z_0)$, *in which the function has a maximum.*
Then the function

$$\mathcal{G}_{z_0} : \Theta \to \mathbb{R} \ \text{with} \ \mathcal{G}_{z_0}(\theta) := \varphi_{z_0}(x(t_0)^T \theta)$$

is said to be a **quality function for polynomial regression**.

Usually, we would take $\mathcal{G}_{z_n}$ to be the likelihood function $\mathcal{G}_{z_n}(\theta) = f_\theta(z_n)$, where f_θ is the density function of Z_n, but also other choices are possible. A family of quality functions is needed to define the global depth (compare Mizera 2002):

Definition 2.2 (Global depth) *Let* $\{\mathcal{G}_{z_0}\}_{z_0 \in \mathcal{Z}}$ *be a family of quality functions for polynomial regression.*
The **global depth** $d_G(\theta, z)$ *of* $\theta = (\theta_1, ..., \theta_q)^T \in \Theta$ *within* $z_1, ..., z_N \in \mathcal{Z}$ *is defined to be the smallest number* m *of observations* $z_{i_1}, ..., z_{i_m}$ *that needs to be removed, such that there is a parameter* $\theta' \neq \theta$ *with*
$\mathcal{G}_{z_j}(\theta') > \mathcal{G}_{z_j}(\theta)$ *for all* $j \in \{1, ..., N\} \backslash \{i_1, ..., i_m\}$.

Roughly speaking, the global depth of a parameter θ with respect to the observations $z_1, ..., z_N \in \mathcal{Z}$ is the number of observations that needs to be removed, such that there is a better parameter for all remaining observations.

Let us define the sign of an observation $z_0 = (y_0, t_0) \in \mathcal{Z}$ to be

$$\text{sig}_\theta(z_0) := \text{sign}(h(z_0) - x(t_0)^T \theta).$$

In most regression setups we have $h(z_n) = y_n$ and thus, $\text{sig}_\theta(z_n)$ is the sign of the residual of the observation z_n. However for regression with exponential distributed dependent observations Y_n we have $h(z_n) = \log(y_n)$ (see Müller 2005), if we take the quality functions to be the likelihood functions. The global depth coincides with the regression depth of Rousseeuw and Hubert (1999), if the dependent observations are appropriately transformed.

It can be shown (see Rousseeuw and Hubert 1999), that the parameter space $\Theta = \mathbb{R}^q$ is divided up into domains (open, connected sets) with constant depth by the hyperplanes

$$H_{z_n} = \{\theta \in \mathbb{R}^q : \text{sig}_\theta(z_n) = 0\}, n = 1, ..., N.$$

For given observations let $\text{Dom}(z)$ be the set of all those domains with constant depth. We define $\tilde{d}_G(G, z) := d_G(\theta, z)$ for $G \in \text{Dom}(z)$ and $\theta \in G$.

The following modification of the global depth leads to better tests, than the global depth itself. Therefor, we modify the global depth on

4

$\text{Border}(z) := \bigcup\limits_{n=1}^{N} H_{z_n}$. The resulting depth function is called the harmonized depth.

Definition 2.3 (Harmonized depth) *The **harmonized depth** of θ within $z = (z_1, ..., z_N)$ is defined to be*

$$d_H(\theta, z) := \min_{G \in \text{Dom}(z), \theta \in \bar{G}} \tilde{d}_G(G, z),$$

where $\bar{G}$ is the closure of G.

This depth has the disadvantage, that the distribution is not known. For the purpose of obtaining tests, we consider the U-statistic with the harmonized depth as the kernel function. The resulting depth function is called the simplicial depth.

Definition 2.4 (Simplicial depth) *The **simplicial depth** of θ within $z = (z_1, ..., z_N)$ is defined to be*

$$d_S(\theta, z) := \binom{N}{q+1}^{-1} \sum_{1 \leq n_1 < ... < n_{q+1} \leq N} d_H(\theta, (z_{n_1}, \ldots, z_{n_{q+1}})).$$

If there are $q + 1$ observations $z_1, ..., z_N \in \mathcal{Z}$, such that $t_1, ..., t_N$ are pairwise different, then there is exactly one bounded domain $S(z) \in \text{Dom}(z)$. The closure of this domain is a simplex. Since the harmonized depth $d_H(\cdot, z)$ is zero on $\text{Border}(z)$, the harmonized depth is the indicator function $d_H(\theta, z) = \mathbb{I}_{S(z)}(\theta)$. Hence, we can write

$$d_S(\theta, z) = \frac{\#\left\{\{n_1, ..., n_{q+1}\} \subset \{1, ..., N\} : d_H\left(\theta, \left(z_{n_1}, ..., z_{n_{q+1}}\right)\right) > 0\right\}}{\binom{N}{q+1}}.$$

This means, that the simplicial depth is the fraction of simplicies, which contains the parameter θ as an interior point.

For testing a hypothesis of the form $H_0 : \theta \in \Theta_0$, the test statistic will be based on $\max\limits_{\theta \in \Theta_0} d_S(\theta, z)$ and the simplicial depth is maximal in a point $\theta \in \mathbb{R}^q \backslash \text{Border}(z)$. The hypothesis H_0 is rejected, if the maximum depth is to small. The asymptotic distribution does not depend on it, how the kernel function is defined on $\text{Border}(z)$, but the test statistic depends on this choice. If the simplicial depth would be based on the global depth as the kernel function, then the depth would be maximal in a point $\theta \in \text{Border}(z)$ and we would have a greater test statistic. This would be a disadvantage, if the aim is to reject the null hypothesis.

For $q + 1$ observations $z_1, ..., z_{q+1} \in \mathcal{Z}$ with pairwise different $t_1, ..., t_{q+1}$ and $\theta \in \mathbb{R}^q \backslash \text{Border}(z)$, we have an important characterization of the harmonized depth. We say, that the observations are alternating, if there is a permutation π with $t_{\pi(1)} < ... < t_{\pi(q+1)}$, such that the sequence $\text{sig}_\theta(z_{\pi(1)}), ..., \text{sig}_\theta(z_{\pi(q+1)})$ is alternating. The harmonized depth is equal to one, iff the observations are alternating. This characterization was used also in Müller (2005) for constructing tests.

2.2 The maximum simplicial depth estimator

The simplicial depth can be used to estimate unknown parameters.

Definition 2.5 (Maximum simplicial depth estimator) *A **maximum simplicial depth estimator** for given observations $z_1, ..., z_N \in \mathcal{Z}$ with respect to a subset $\mathbb{K} \subset \mathbb{R}^q$ is defined to be a parameter*

$$\hat{\theta}_S \in \overline{\arg \max_{\theta \in \mathbb{K}} d_S(\theta, z)}.$$

We usually choose $\mathbb{K} = \mathbb{R}^q$, but if we assume that the true parameter belongs to a known subset of $\mathbb{R}^q$, then we can take $\mathbb{K}$ to be this subset. In particular for testing $H_0 : \theta \in \Theta_0$ we choose $\mathbb{K} = \Theta_0$.

The maximum simplicial depth estimator is not unique. If $\mathbb{K}$ is an affine subspace of $\mathbb{R}^q$ or a polyhedron, then the closure of the set of all parameters $\theta \in \mathbb{K}$ that maximize $d_S(\cdot, z)|_{\mathbb{K}}$ is a union of polytopes. Let $\mathbb{P}$ be the set of these polytopes. We will calculate the verticies $\text{ext}(P)$ of each polytope $P \in \mathbb{P}$. Then we have

$$\overline{\arg \max_{\theta \in \mathbb{K}} d_S(\theta, z)} = \bigcup_{P \in \mathbb{P}} \text{conv}(\text{ext}(P)), \tag{1}$$

where $P = \text{conv}(\text{ext}(P))$ is the set of all convex combinations of verticies from P. If we assume, that the true probability measure belongs to $\{P_\theta : \theta \in \mathbb{K}\}$, then we can choose $\hat{\theta}_S \in \bigcup_{P \in \mathbb{P}} P$ as an estimation for the true parameter.

3 Computation

In this section, we compute maximum simplicial depth estimators, defined in 2.5, where $\mathbb{K}$ is an affine subspace of $\mathbb{R}^q$, a polyhedron or just a single point. Therefore, let be $z_1, ..., z_N \in \mathcal{Z}$, such that $t_1, ..., t_N$ are pairwise different and

suppose, that there is no polynomial of degree $q - 1$, that contains $q + 1$ observations.

3.1 Computing the simplicial depth of a given parameter

For computing maximum simplicial depth estimators, we need an algorithm for calculating the simplicial depth of a given parameter value $\theta \in \mathbb{R}^q \backslash \mathrm{Border}(z)$. The simplicial depth doesn't depend on the order of the observations, so we may assume, that the observations are already ordered, this means $t_1 < ... < t_N$.

Since the harmonized depth for $q + 1$ observations is equal to one, iff the signs of the observations are alternating, the absolute simplicial depth $\binom{N}{q+1} d_S(\theta, z)$ is equal to the number of ordered $q + 1$-tuples $(n_1, \ldots, n_{q+1}) \in \{1, \ldots, N\}^{q+1}$ for which the sequence $(\mathrm{sig}_\theta(z_{n_1}), ..., \mathrm{sig}_\theta(z_{n_{q+1}}))$ is alternating.

Let l_r be the number of runs within $(\mathrm{sig}_\theta(z_1), ..., \mathrm{sig}_\theta(z_N))$. For $j \leq l_r$ take R_j to be the set of all $n \in \{1, ..., N\}$, such that z_n belongs to the j-th run and let $r_j := \#R_j$. For $j > l_r$ let $r_j = 0$.

Then we have

$$
\binom{N}{q+1} d_S(\theta, z) = \# \bigcup_{\substack{1 \leq j_1 < ... < j_{q+1} \leq l_r \\ j_{k+1} - j_k \text{ is odd}}} R_{j_1} \times ... \times R_{j_{q+1}}
$$

$$
= \sum_{\substack{1 \leq j_1 < ... < j_{q+1} \\ j_{k+1} - j_k \text{ is odd}}} r_{j_1} \cdot ... \cdot r_{j_{q+1}}
$$

$$
= \sum_{j_1=1}^{l_r} r_{j_1} \Big(\sum_{j_2 \in j_1 + 1 + 2\mathbb{N}_0} r_{j_2} \big(... r_{j_q} \big(\sum_{j_{q+1} \in j_q + 1 + 2\mathbb{N}_0} r_{j_{q+1}} \big) ... \big) \Big)
$$

$$
= \sum_{j=1}^{l_r} h_{q+1}(j),
$$

where h_k is recursively defined by $h_1(j) := r(j)$ and

$$
h_k(j) := r_j \sum_{i \in j + 1 + 2\mathbb{N}_0} h_{k-1}(i) \text{ for } k = 2, ..., q + 1 \text{ and } j \geq 1.
$$

The number $h_k(j)$ is the number of all ordered k-tuples of observations with alternating signs, for which the smallest index belongs to the j-th run.

Let

$$g_k(j) := \sum_{i \in j+1+2\mathbb{N}_0} h_k(i) \text{ for } k = 1, ..., q \text{ and } j \geq 1.$$

Then we have the recursion formulas

$$g_k(j) = h_k(j + 1) + g_k(j + 2) \text{ and}$$
$$h_k(j) = r_j \, g_{k-1}(j).$$

Now let sig be a shorthand for the sequence $(\mathrm{sig}_\theta(z_1), ..., \mathrm{sig}_\theta(z_N))$ and suppose that the observations $z_1, ..., z_N$ are ordered. Then the following algorithm calculates the simplicial depth:

```
"sim.depth.sig" ← function(q, sig){
    N ← length(sig)

    # — Calculation of the run lengths —
    r ← 0 ∈ ℝ^N
    cur.run ← 1
    cur.sig ← sig_1
    for j from 1 to N do {
        if(sig_j ≠ cur.sig){
            cur.run ← cur.run + 1
            cur.sig ← sig_j
        }
        r_{cur.run} ← r_{cur.run} + 1
    }

    # — Calculation of the depth —
    l_r ← cur.run
    h_1 ← r
    if(l_r ≥ q + 1){
        for k from 2 to (q + 1) do {
            for j from (l_r − 1) to 1 step −1 do {
                g_{k−1}(j) ← h_{k−1}(j + 1) + g_{k−1}(j + 2)
                h_k(j) ← r_j g_{k−1}(j)
            }
        }
    }
    absolute.depth ← Σ_{j=1}^{l_r} h_{q+1}(j)
    return absolute.depth
}
```

3.2 Assumptions and notations for computing the maximum simplicial depth estimators

In this subsection, we introduce the notations, that are used within the algorithm, which computes all maximum simplicial depth estimators with respect to an affine subspace $I\!K$.

Because of (1), it suffices to calculate the verticies of the polytopes $P \in I\!\!P$.

Let $D := \dim I\!K$ and let J_D be the set of all ordered D-tuples $(j_1, ..., j_D) \in \{1, ..., N\}^D$.

For brevity let us write $X(j) := \begin{pmatrix} x(t_{j_1})^T \\ : \\ x(t_{j_D})^T \end{pmatrix}$ and $h(j) := \begin{pmatrix} h(z_{j_1}) \\ : \\ h(z_{j_D}) \end{pmatrix}$, where $j \in J_D$. The affine subspace $I\!K$ has a representation

$$I\!K = p + \mathbb{R}r_1 + ... + \mathbb{R}r_D$$

with $p \in \mathbb{R}^q$ and linearly independent $r_1, ..., r_D \in \mathbb{R}^q$. Let $R := (r_1, ..., r_D)$ and suppose, that for each subset of D observations $z_{j_1}, ..., z_{j_D}$ from $z_1, ..., z_N$ the matrix $X(j) R$ can be inverted.

For each $P \in I\!\!P$ there is a $G \in \text{Dom}(z)$, such that $P = \bar{G} \cap I\!K$. The verticies of each polytope $\bar{G} \cap I\!K$ with $G \in \text{Dom}(z)$ are points of intersection of D hyperplanes from $H_{z_1}, ..., H_{z_N}$ with $I\!K$. For $j \in J_D$ let $\tilde{\theta}_j \in \mathbb{R}^q$, such that

$$\{\tilde{\theta}_j\} = \bigcap_{i=1}^{D} H_{z_{j_i}} \cap I\!K \ \ (\text{ see } 3.3).$$

With the algorithm, we find all parameters from $\{\tilde{\theta}_j : j \in J_D\}$, which are verticies of a polytope $P \in I\!\!P$ with maximum depth. The algorithm calculates the maximum depth with the formula

$$\max_{\theta \in I\!K} d_S(\theta, z) = \max_{j \in J} \max_{G \in \text{Dom}(z) : \tilde{\theta}_j \in \bar{G}} d_S(G, z).$$

3.3 The algorithm for computing the maximum simplicial depth estimators

The algorithm for computing the verticies of the polytopes $P \in I\!\!P$ with maximum depth depends on $t := (t_1, ..., t_N)$, $(h(z_1), ..., h(z_N))$, p and R. By using the same notations as in the text, the following algorithm calculates the required verticies. For each vertex it calculates an ID (identification vector) for

the adjacent domain with maximum depth. As an ID we use the signs, that the observations have with respect to a parameter from this domain. Furthermore, it calculates the test statistic for the one-sample test, described in the next section.

1    $\pi \leftarrow \textbf{order}(t)$
2    $t \leftarrow (t_{\pi(1)}, ..., t_{\pi(N)})^T$
3    $h \leftarrow (h(z_{\pi(1)}), ..., h(z_{\pi(N)}))^T$
4    $X \leftarrow \begin{pmatrix} 1 & t_1 & ... & t_1^{q-1} \\ : & : & : & : \\ 1 & t_N & ... & t_N^{q-1} \end{pmatrix}$
5    $\hat{\theta}_{ls} \leftarrow p + R\,\textbf{lsfit}(X\,R, h - X\,p)$
6   lower.bound $\leftarrow$ sim.depth.sig$(q, \textbf{sign}(h - X\,\hat{\theta}_{ls}))$
7   param $\leftarrow \emptyset$
8    $\textbf{for}(j \in J_D)\{$
9      $\lambda_j \leftarrow (X(j)\,R)^{-1}\,(h(j) - X(j)\,p)$
10     $\tilde{\theta}_j \leftarrow p + R\,\lambda_j$
11   sig $\leftarrow \textbf{sign}(h - X\,\tilde{\theta}_j)$
12     $\textbf{for}(s \in \{-1, 1\}^D)\{$
13       $\textbf{for}(m \in \{1, ..., D\})\{\text{sig}_{j_m} \leftarrow s_m\}$
14   depth $\leftarrow$ sim.depth.sig(q, sig)
15       $\textbf{if}(\text{depth} > \text{lower.bound})\{$
16   param $\leftarrow \emptyset$
17   lower.bound $\leftarrow$ depth$\}$
18       $\textbf{if}(\text{depth} == \text{lower.bound})\{$
19   param $\leftarrow$ param $\cup \{(\tilde{\theta}_j, \text{sig})\}$
20       $\}$
21     $\}$
22    $\}$
23    $\textbf{return}$ list(best.param = param, test.statistic = $N\left(\frac{\text{lower.bound}}{\binom{N}{q+1}} - \frac{1}{2^q}\right)$)

At first, we sort the observations, such that $t_1 < ... < t_N$. In line 5 and 6, we calculate a lower bound of the simplicial depth, namely the simplicial depth of the LS estimation of θ.

In the lines 8-22, we consider all verticies one after another in a suitable order. As mentioned in 3.2, each subset of D observations $z_{j_1}, ..., z_{j_q}$ from $z_1, ..., z_N$ corresponds to a vertex $\tilde{\theta}_j$.

In line 10, we calculate this vertex. Let us proof the formulas, mentioned in line 9 and 10. Since $\tilde{\theta}_j \in H_{z_{j_i}}$ for $i = 1, ..., D$, we have $X(j)\,\tilde{\theta}_j = h(j)$.

Since $\tilde{\theta}_j \in I\!\!K$, there exists a $\lambda_j \in \mathbb{R}^D$, such that $\tilde{\theta}_j = p + R\,\lambda_j$ and we have $\lambda_j = (X(j)\,R)^{-1}(h(j) - X(j)\,p)$.

Later on, in the lines 12-21, we calculate the depths of the 2^D adjacent domains. In order to improve the algorithm, it could be modified, such that it stops the calculation of the depths of the 2^D adjacent domains as soon as it is clear, that none of them has maximum depth.

We put each parameter with maximum depth and the sign vector of the adjacent domain with maximum depth into the list "param".

It remains to say, how to calculate the depths of the adjacent domains. The depth of an adjacent domain G depends only on the sequence $\mathrm{sig}_\theta(z_1), ..., \mathrm{sig}_\theta(z_N)$, where $\theta \in G$. For $n \in \{1, ..., N\} \backslash \{j_1, ..., j_D\}$ we have $\mathrm{sig}_\theta(z_n) = \mathrm{sig}_{\tilde{\theta}_j}(z_n)$ and for each $s \in \{-1, 1\}^D$, there is exactly one adjacent domain with $(\mathrm{sig}_\theta(z_{j_1}), ..., \mathrm{sig}_\theta(z_{j_D})) = s$.

At last, we create a list with the results in line 23. The parameters with the same sign vector create the same polytope.

3.4   Calculating the maximum depth for a polyhedron

If $I\!\!K$ is a polyhedron within an affine D-dimensional subspace of $\mathbb{R}^q$, then we have to modify the algorithm for calculating all parameters $\theta \in \arg\max\limits_{\theta \in I\!\!K} d_S(\theta, z)$. In particular, we need another approach for calculating the lower bound given by the depth of the least square fit in line 6, because it could happen, that the least square fit doesn't belong to $I\!\!K$. The easiest way is to choose 0 as the lower bound.

The boundary of $I\!\!K$ is a union of hyperplanes. We put them to the hyperplanes $H_{z_1}, ..., H_{z_N}$ and obtain a set $\mathcal{H}$ of hyperplanes. If we take D hyperplanes from $\mathcal{H}$, then they have a point of intersection. For each point of intersection, that belongs to $I\!\!K$, we calculate the depths of the adjacent domains, that have a nonempty intersection with $I\!\!K$. In this way, we find the verticies of the set of all parameters with maximum depth within $I\!\!K$.

3.5   The best of all deepest parameters

In general the set $\overline{\arg\max\limits_{\theta \in I\!\!K} d_S(\theta, z)} = \bigcup\limits_{P \in I\!\!P} \mathrm{conv}(\mathrm{ext}(P))$ of all deepest parameters given by (1) consists of more than one parameter. Then the problem is how to choose a parameter from this set. We propose here two methods based

on l_1 and l_2 minimization to choose a best deepest parameter.

For $P \in \mathbb{P}$, let us write $ext\,(P) = \left\{\theta_{P,1}, ..., \theta_{P,N_p}\right\}$. Let Λ_P be the set of all $\lambda \in [0,1]^{N_p}$ with $\sum_{i=1}^{N_p} \lambda_i = 1$. For each $\theta \in P$ there exists a $\lambda \in \Lambda_P$ such that

$$\theta = \theta_{P,1}\lambda_1 + ... + \theta_{P,N_p}\lambda_{N_p}.$$

With $\theta_P := \left(\theta_{P,1}, ..., \theta_{P,N_p}\right)$ we have $\theta = \theta_P \lambda$, where $\lambda = \left(\lambda_1, ..., \lambda_{N_p}\right)^T$. We define the optimal parameter as that $\widehat{\theta}_{BD} = \theta_P\hat{\lambda} \in \cup_{P \in \mathbb{P}} P$ with

$$\sum_{n=1}^{N} \rho\left(y_n - x\,(t_n)^T\,\theta_P\hat{\lambda}\right) = \min_{\lambda \in \Lambda_P,\, P \in \mathbb{P}} \sum_{n=1}^{N} \rho\left(y_n - x\,(t_n)^T\,\theta_P\lambda\right),$$

where $\rho\,(x) = |x|$ or x^2. In examples, we find the solution by calculating $M > 20000$ randomly generated vectors $\lambda \in \Lambda_P$. The parameter $\widehat{\theta}_{BD}$ is called the **b**est of all **d**eepest parameters.

4 Tests

4.1 One-sample tests

For the one-sample problem, the distribution free asymptotic α-level tests for testing all hypotheses of the form $H_0 : \theta \in \Theta_0$, where Θ_0 is a subset of the parameter space, are based on the simplicial depth.

We assume, that there is a known $p \in [0,1]$ such that for all $\theta \in \Theta$ we have:

- $P_\theta(\mathrm{sig}_\theta(Z_1) \geq 0|T_1) \equiv p$ a.s., $\hspace{3cm}$ (2)
- $P_\theta(\mathrm{sig}_\theta(Z_1) = 0|T_1) = 0$ a.s.,
- $P_\theta(T_1 = t) = 0$ for all $t \in \mathbb{R}$.

The asymptotic distribution depends on it, if $p = \frac{1}{2}$ or $p \neq \frac{1}{2}$. The most interesting case is $p = \frac{1}{2}$. For the other one see Müller (2005). In both cases, the test statistic is based on $\sup_{\theta \in \Theta_0} d_S(\theta, (z_1, ..., z_N))$ and can be computed with the algorithm from 3.3 resp. 3.4, if Θ_0 is an affine subspace or a polyhedron.

In particular, for $p = \frac{1}{2}$ the test statistic is given by

$$T(z_1, ..., z_N) := N\left(\sup_{\theta \in \Theta_0} d_S(\theta, (z_1, ..., z_N)) - \frac{1}{2^q}\right).$$

Its asymptotic distribution was derived for linear and quadratic regression in Müller (2005) and for general ploynomial regression in Wellmann et al. (2006). It is given as follows:

If q is even, then

$$N\Big(d_S(Z_1,\ldots,Z_N) - \frac{1}{2^q}\Big) \xrightarrow{\mathcal{L}} \sum_{l=0}^{\infty} \lambda_{2l+1}\Big(X_l^2 + Y_l^2 - 2\Big), \qquad (3)$$

if q is odd, then

$$N\Big(d_S(Z_1,\ldots,Z_N) - \frac{1}{2^q}\Big) \xrightarrow{\mathcal{L}} \lambda_0\Big(W^2 - 1\Big) + \sum_{l=1}^{\infty} \lambda_{2l}\Big(X_l^2 + Y_l^2 - 2\Big), \qquad (4)$$

where

- $W, X_0, Y_0, X_1, Y_1, \ldots$ are i.i.d. standard normal distributed random variables.
- $\lambda_0 = -\frac{(q+1)}{2^{q+1}}$
- $\lambda_j = \sum_{\substack{k \in \{0,\ldots,q-1\} \\ k \text{ odd}}} \frac{(q+1)!}{2^{q-k}(q-1-k)!}\Big(-j^2\pi^2\Big)^{-\frac{(k+1)}{2}}$ for $j \in \mathbb{N}$.

The null hypothesis H_0 is rejected, if $T(z_1,\ldots,z_N)$ is less than the α-quantile of the asymptotic distribution, mentioned in (3) and (4).

A simple possibility for estimating the quantiles is the generation of random numbers of the distribution. The quantiles given in the Table 1 where calculated by computing 10 000 random numbers of the distribution (only the first 200 summands). The calculation of the quantiles was repeated 500 times. The means of these quantiles are given in the table. The 99.5% confidence band is less then ± 0.004 for each estimated quantile and it is less then ± 0.001 for $\alpha \geq 6\%$.

4.2   Two-sample tests

Let $\mathbb{K}$ be a subset of the parameter space $\Theta = \mathbb{R}^q$. In this section, we present a distribution free, asymptotic α-level test for testing the hypothesis, that independent observations from two populations can be described by the same polynomial regression function with parameter in $\mathbb{K}$.

For $i = 1, 2$ take θ^i to be the unknown, true parameter for the observations $z^i := (z_{i,1}, \ldots, z_{i,N_i}) \in \mathcal{Z}^{N_i}$ from the i-th sample. For each population, we make the same assumptions, as for the one sample test. In particular, we assume, that (2) holds.

Table 1

Means of the simulated quantiles for different regression models

α-quantile	linear	quadratic	cubic	
	$q = 2$	$q = 3$	$q = 4$	$q = 5$
0.5%	-2.617	-1.942	-1.312	-0.840
1.0%	-2.195	-1.633	-1.112	-0.716
1.5%	-1.949	-1.454	-0.994	-0.643
2.0%	-1.775	-1.327	-0.911	-0.591
2.5%	-1.639	-1.228	-0.847	-0.550
3.0%	-1.529	-1.148	-0.793	-0.517
4.0%	-1.354	-1.021	-0.709	-0.464
5.0%	-1.218	-0.923	-0.644	-0.422
6.0%	-1.107	-0.843	-0.591	-0.388
7.0%	-1.013	-0.775	-0.545	-0.359
8.0%	-0.932	-0.716	-0.506	-0.334
9.0%	-0.861	-0.665	-0.471	-0.312
10.0%	-0.797	-0.618	-0.439	-0.292
15.0%	-0.550	-0.438	-0.317	-0.213
20.0%	-0.375	-0.310	-0.228	-0.156
25.0%	-0.240	-0.209	-0.158	-0.110
30.0%	-0.129	-0.125	-0.010	-0.071
40.0%	0.046	0.010	-0.000	-0.008
50.0%	0.182	0.119	0.007	0.045
60.0%	0.293	0.212	0.143	0.092
70.0%	0.388	0.296	0.208	0.137
80.0%	0.473	0.378	0.272	0.182
90.0%	0.555	0.466	0.343	0.234
95.0%	0.601	0.520	0.390	0.269
99.5%	0.668	0.611	0.473	0.334

We don't reject the hypothesis $H_0 : \theta^1 = \theta^2 \in I\!K$ with respect to a significance level α, if there is a $\theta \in I\!K$ such that neither the hypothesis that θ is the true parameter for the first population, nor the hypothesis that θ is the true parameter for the second population can be rejected with respect to a significance level $\frac{\alpha}{2}$. Hence, we use the test statistic

$$T(z^1, z^2) := \max_{\theta \in I\!K} \Phi_\theta(z^1, z^2),$$

where

$$\Phi_\theta(z^1, z^2) := \min\left(N_1(d_S(\theta, z^1) - \frac{1}{2^q}), N_2(d_S(\theta, z^2) - \frac{1}{2^q})\right)$$

and we reject $H_0 : \theta^1 = \theta^2 \in I\!K$, if $T(z^1, z^2)$ is less than the $\frac{\alpha}{2}$-quantil of the distribution, mentioned in (3), respectively (4).

Now, we prove that this is an asymptotic α-level test. Let $\varphi_N^{\tilde\theta} : \mathbb{R}^N \to \{0, 1\}$ be the asymptotic one-sample $\frac{\alpha}{2}$-level test for testing the hypothesis $H_0 : \theta = \tilde\theta$. We want to test the hypothesis $H_0 : (\theta^1, \theta^2) \in \Theta_0 := \{(\theta', \theta'') \in I\!K^2 : \theta' = \theta''\}$ as described above. Let $Z^1 = (Z_{11}, ..., Z_{1N_1})$, $Z^2 = (Z_{21}, ..., Z_{2N_2})$ and let $\psi_{N_1, N_2}(z^1, z^2)$ be the test function.

For all $\theta_0 = (\theta', \theta') \in \Theta_0$ we have

$$P_{\theta_0}(\exists \tilde\theta \in I\!K : \varphi_{N_1}^{\tilde\theta}(Z^1) = 0 \text{ and } \varphi_{N_2}^{\tilde\theta}(Z^2) = 0)$$
$$\geq P_{\theta_0}(\varphi_{N_1}^{\theta'}(Z^1) = 0 \text{ and } \varphi_{N_2}^{\theta'}(Z^2) = 0)$$
$$\overset{\text{indep.}}{=} P_{\theta_0}(\varphi_{N_1}^{\theta'}(Z^1) = 0) \, P_{\theta_0}(\varphi_{N_2}^{\theta'}(Z^2) = 0).$$

Hence,

$$\lim \sup_{N_1, N_2 \to \infty} P_{\theta_0}(\psi_{N_1, N_2}(Z^1, Z^2) = 1)$$
$$= 1 - \lim \inf_{N_1, N_2 \to \infty} P_{\theta_0}(\psi_{N_1, N_2}(Z^1, Z^2) = 0)$$
$$\overset{\text{Def.}}{=} 1 - \lim \inf_{N_1, N_2 \to \infty} P_{\theta_0}(\exists \tilde\theta \in I\!K : \varphi_{N_1}^{\tilde\theta}(Z^1) = 0 \text{ and } \varphi_{N_2}^{\tilde\theta}(Z^2) = 0)$$
$$\leq 1 - \lim \inf_{N_1, N_2 \to \infty} P_{\theta_0}(\varphi_{N_1}^{\theta'}(Z^1) = 0) \, P_{\theta_0}(\varphi_{N_2}^{\theta'}(Z^2) = 0)$$
$$\leq 1 - \left(1 - \frac{\alpha}{2}\right)\left(1 - \frac{\alpha}{2}\right) < \alpha.$$

15

4.3 Computing the two-sample test statistic

Suppose, that $I\!K$ is an affine subspace. The subspace $I\!K$ is divided up into domains by the hyperplanes $H_{z_{1,1}}, ..., H_{z_{1,N_1}}, H_{z_{2,1}}, ..., H_{z_{2,N_2}}$. In each domain, both depth functions are constant. This means that $d_S(\cdot, z^i)|_{G \cap I\!K}$ is constant for $i = 1, 2$ and all $G \in \mathrm{Dom}(z^1, z^2)$.

Hence, we put both populations together and we calculate the verticies of each polytope $\bar{G} \cap I\!K$ with $G \in \mathrm{Dom}(z^1, z^2)$. Let $\{\tilde{\theta}_j : j \in J\}$ be the set of all verticies.

Then it remains to calculate

$$T(z^1, z^2) = \max_{j \in J} \max_{G \in \mathrm{Dom}(z^1, z^2): \tilde{\theta}_j \in \bar{G}} \Phi_G(z^1, z^2).$$

For this purpose, the algorithm in 3.3 can easily be modified.

5 Comparison with other estimates

In this section we give a brief summary of the other selected estimates, and discuss and compare used estimates in examples from astronomy. For calculating the simplicial depth, we choose $h(z_n) = y_n$ within the examples.

5.1 Other regarded estimators

In order to compare maximum simplicial depth estimators with some other selected estimates, we give a brief summary of these techniques, namely ordinary least square (OLS), l_1-, M-, MM- and CAT- estimators. All regarded estimators are implemented in S-Plus (see Marazzi, 1997).

The **OLS** (l_2) **estimate**

$$\widehat{\theta}_{l_2} \in \arg\min_{\theta \in \Theta} \sum_{n=1}^{N} \left(y_n - x\left(t_n \right)^T \theta \right)^2$$

is calculated with the S-Plus function `lm`.

The **least absolute deviation** (l_1) **estimate**

$$\widehat{\theta}_{l_1} \in \arg\min_{\theta \in \Theta} \sum_{n=1}^{N} \left| y_n - x\left(t_n \right)^T \theta \right|.$$

is computed by the function `l1fit`.

Robust M- estimators $\widehat{\theta}_M$ of θ are defined as

$$\widehat{\theta}_M \in \arg\min_{\theta \in \Theta} \sum_{n=1}^{N} \rho\left(\frac{y_n - x\left(t_n\right)^T \theta}{\hat{s}}\right),$$

where $\rho\left(\cdot\right)$ is score function of the residuals with continuous and bounded derivative ψ and $\hat{s}$ is a robust scale estimator for the residuals. In the examples we use $-\rho$ as the likelihood function of Cauchy distribution (see Jurečková & Sen, 1996) as a function which down weights outliers because of its convergence to zero when $y \longmapsto \pm\infty$. This is in particular a suitable estimator, if we expect outliers in y-direction. It is calculated by the function `rreg`, `method=wt.cauchy`.

To deal with outliers in both y-direction and $x\left(t\right)$-space, we use MM-estimates (see Yohai 1987) as a combination of high breakdown value estimation and M-estimation. It has both the high breakdown property and higher statistical efficiency than S-estimation and the least trimmed squares (LTS) estimation. The regarded robust MM-estimator is computed by the S-Plus function `lmRobMM`.

The **LTS estimate** $\widehat{\theta}_{LTS}$ of θ is

$$\widehat{\theta}_{LTS} \in \arg\min_{\theta \in \Theta} \sum_{n=1}^{h} r_{(n)}^2\left(\theta\right),$$

where $r_{(1)}^2\left(\theta\right) \leq r_{(2)}^2\left(\theta\right) \leq ... \leq r_{(N)}^2\left(\theta\right)$ are the ordered squared residuals $r_n^2\left(\theta\right) = \left(y_n - x\left(t_n\right)^T \theta\right)^2$, $n = 1, 2, ..., N$ (see Rousseeuw, 1984). It is calculated by the function `ltsreg` with default $h = \lfloor max((N + q + 1)/2, 0.9N)\rfloor$.

Another outlier robust estimator for linear regression is the **catline** (CAT comes from Cuts All Thirds), defined in Hubert and Rousseeuw (1999). If $q = 2$, then we can take the catline as an estimator for the true parameter. In general, the catline is not the maximum depth estimator for the global depth, but it is shown by Hubert and Rousseeuw (1999), that the catline maximizes it in special cases.

While the first five S-Plus functions come from the basic S-Plus 6.2 package, the CAT-estimator `catline.s` is from Robust Statistics web page (http://wis.kuleuven.be/stat/robust.html).

Data used in this application are from the Hertzsprung-Russell Diagram of the Star Cluster CYG OB1, which contains 46 stars in the direction of Cygnus, from Rousseeuw and Leroy (1987). The first variable is the logarithm of the effective temperature at the surface of the star (t_n) and the second one is the logarithm of its light intensity (y_n). In Figure 1, which is the scatterplot of these data points, two groups of points are seen: the majority which tend to follow a steep band and four stars in the upper left corner (called giants).

In the linear regression model with $q = 2$, the best parameter in the sense of l_1 and l_2 minimization (see Subsection 3.5) within the single polytope with maximum depth is $\widehat{\theta}_{BD} = (-4.698, 2.180824)^T$, which is one vertex of this four-vertex polytope with maximum simplicial depth 0.134 (see Figure 1).

In the quadratic regression model with $q = 3$ and $\theta_3 = 0$, the best parameter in the sense of l_1 and l_2 minimization within the single polytope with maximum depth is $\widehat{\theta}_{BD} = (-8.206, 2.990, 0)^T$, which is one vertex of this three-vertex polytope with maximum simplicial depth 0.134 (see Figure 1). It is interesting to see the comparison of the linear models with $q = 2$ versus $q = 3$ with $\theta_3 = 0$, where the estimators and also the polytopes with maximum depth are different.
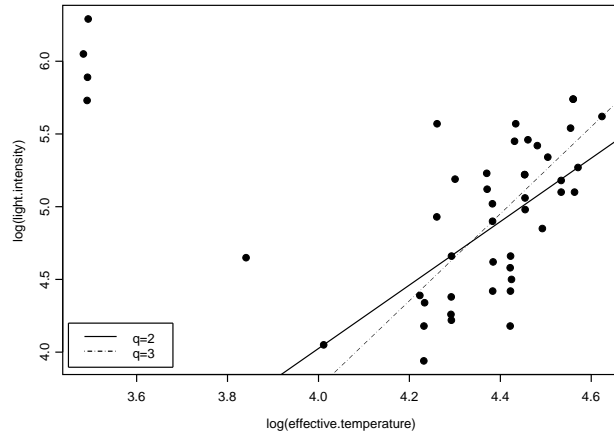
Fig. 1. The best deepest lines ($q = 2$ and $q = 3$)

The other regarded estimates are as follows: $\widehat{\theta}_{l_2} = (6.846, -0.427)^T$, $\widehat{\theta}_{l_1} = (8.158, -0.695)^T$, and $\widehat{\theta}_M = (6.895, -0.435)^T$, which are not robust in the sense of $x(t)$-space outliers, and $\widehat{\theta}_{MM} = (-7.966, 2.924)^T$, $\widehat{\theta}_{LTS} = (-7.550, 2.832)^T$, and $\widehat{\theta}_{CAT} = (-6.183, 2.529)^T$ which are robust, but not deepest estimates (Figure 2 and 3). From Figure 2 and 3, it can be seen that the LTS- and MM-estimate are close to the best deepest line for $q = 3$,

while the cat line is more close to the best deepest line for $q = 2$.
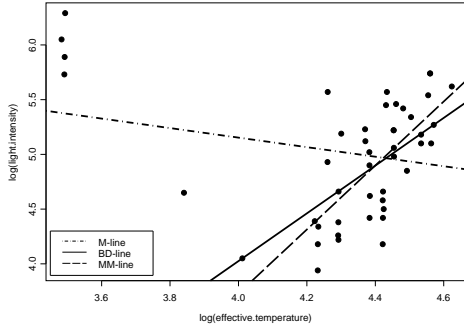

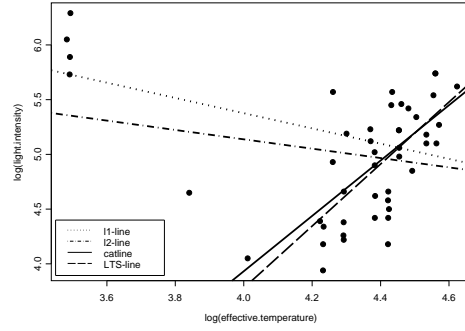Fig. 2. The best deepest line ($q = 2$), M- and MM-line

Fig. 3. Catline, l1-, l2- and LTS-line

In the model for quadratic regression (with $q = 3$ and $\theta_3 \neq 0$), the best parameter in the sense of l_1 and l_2 minimization is $\widehat{\theta}_{BD} = (88.905, -41.734, 5.143)^T$, which is one vertex of the six-vertex polytope with maximum simplicial depth 0.137 (Figure 4). The other regarded estimates are as follows:
$\widehat{\theta}_{l_2} = (82.028, -38.096, 4.671)^T$, and $\widehat{\theta}_{l_1} = (86.412, -40.252, 4.932)^T$, $\widehat{\theta}_{LTS} = (90.222, -42.200, 5.177)^T$, and $\widehat{\theta}_{MM} = (87.021, -40.626, 4.987)^T$ (Figure 5 and 6).
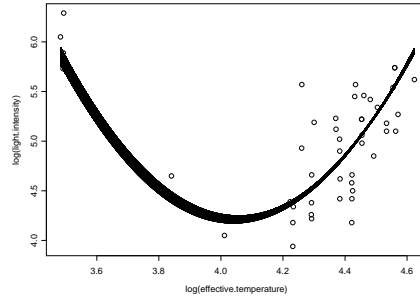
Fig. 4. All deepest quadratic lines
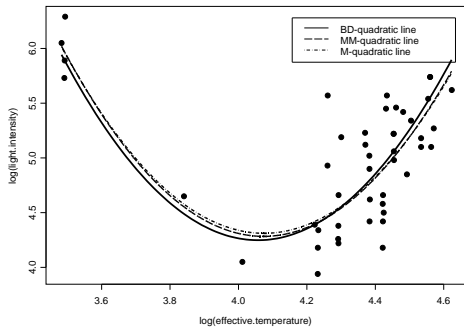
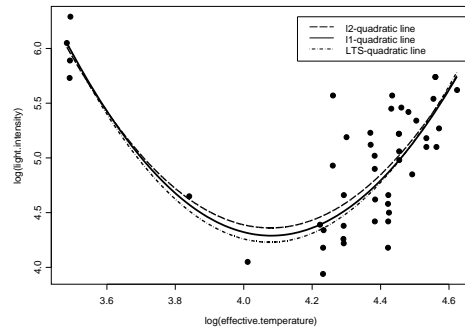Fig. 5. The best deepest quadratic line ($q = 3$), M- and MM-line

Fig. 6. l1-, l2- and LTS-quadratic line

19

6 Tests for a problem in shape analysis

The North American Sunfish "pumpkinseed" (*Lepomis gibbosus*) was introduced to European waters about 100 years ago. In this section we want to find out, how the relative head length of those fishes depends on their size. Furthermore, we investigate, if this dependency can be described by the same polynomial regression function in different European and Canadian populations.

In Canada, 85 specimens were collected in 2003 from the *River Otonabee* (**oto**) and 117 specimens from the *Looncall Lake* (**loon**). A total of 170 specimens were taken from *thermal lagoon Topla struga* near Catez (**cat**), Slovenia. A total of 162 specimens were taken from *Tanyards fisheries pond* near Brighton (**eng**). All specimens were preserved in 4% formaldehyde.

Nineteen landmarks (see Figure 7) were identified on digital photographs using IMPORPRO 3.2 software (Tomeček et.al, 2005 and Katina 2004). The standard length (SL) of a fish is defined as the distance between landmark 10 (anterior tip of the upper jaw) and landmark 11 (caudal fin base). The line, which contains landmark 10 and landmark 11 is called the baseline. The distance between landmark 10 and the orthogonal projection 17' of landmark 17 (occipitum of dorsal outline) onto the baseline is called the head length in this paper. The relative head length is the head length, divided by the standard length.
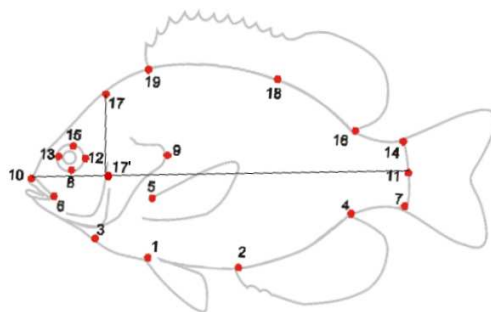

Fig. 7. Landmarks and baseline

The original data are discrete, due to rounding errors. To make them continuous, we add a small uniformly distributed random number to each observation, such that we would obtain the original data by rounding.

At first, we test within the model for quadratic regression the hypothesis, that the relationship between the relative head length and the standard length can be described by a linear function. This hypothesis has to be rejected only

for the population **eng** with respect to a significance level of 10%. The best deepest linear function in the sense of Section 3.5 for **eng** is shown in Figure 8. The test statistics for **oto**, **loon**, **cat** and **eng** are 0.370, 0.359, 0.306 and -0.769 (in this order), providing p-values > 70% for the first three populations and a p-value less then 8% for the population **eng** (see Table 1).

Normality of ordinary least squares residuals, in both linear and quadratic model, was rejected for all four fish populations, **oto**, **loon**, **cat** and **eng** (p-value< 0.001) by the χ^2 goodness-of-fit test. Although the data have no normal distribution, we apply the F-test in order to compare the results with this test. The F-test provides p-values as follows 0.794, 0.031, 0.225 and < 0.0001. Hence, we reject the null hypothesis about linearity for populations **loon** and **eng** with the F-test. In **loon** population it is due to outliers, since the depth test does not reject the linearity. However for the population **eng** it can be concluded that the linear relationship is not valid since both tests reject the hypothesis of linearity.
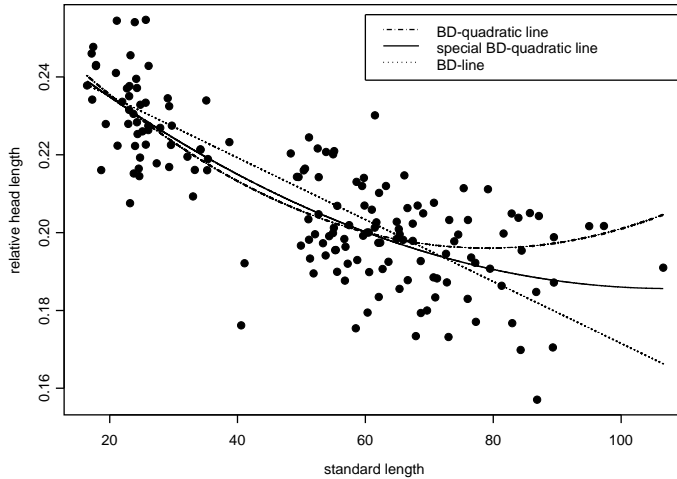

Fig. 8. The population **eng**

The best deepest quadratic function for **eng** is also seen in Figure 8. If it would be the true regression function, then the relative head length decreases when small fishes are growing, but when they become adult, then the relative head length increases. We know no biological reason for this and so we test the hypothesis, that the true regression function is convex and monotonically decreasing all the time. This means, that the derivative of the true regression function g_θ is negative also for big fishes with a standard length of 107. Hence, we have to test $H_0 : g'_\theta(107) \leq 0$ and $g''_\theta \geq 0$, which is equivalent to $H_0 : \theta \in \Theta_0$ with

$$\Theta_0 := \{\tilde{\theta} \in \mathbb{R}^3 : \tilde{\theta}_2 + 214\,\tilde{\theta}_3 \leq 0 \text{ and } \tilde{\theta}_3 \geq 0\}.$$

21

The boundary of this set consists on hyperplanes and thus we can calculate the test statistic as described in Section 3.4. Note, that it is not so easy to test this hypothesis in the classical way. The test statistic is 0.261, which is more than the 60% quantile of the asymptotic distribution and thus we may assume, that the true regression function is monotonically decreasing. The best deepest decreasing quadratic function, called special BD-quadratic line, is shown in Figure 8 and has the parameter $(0.259919, -0.001380, 0.000006)^T$.

Now we investigate, whether the Canadian populations **oto** and **loon** can be described by the same regression line. This can be tested within a model for linear regression, but we could test it as well in a model for polynomial regression of arbitrary degree. An interesting problem is to find out, which test is the best one. The test statistic for the two-sample test is -0.049 for $q = 2$ (and -0.02 for $q = 3$), which is in both cases more than the 30% quantile of the asymptotic distribution (see Table 1). Hence, we have no rejection. This means, that we may assume that both regression lines are equal. Indeed, the deepest lines for **loon** and **oto** drawn into Figure 9 are rather similar. Into this diagram, we also draw a dashed line, that maximizes $\Phi_\theta(z^1, z^2)$, which we call the H_0-line.
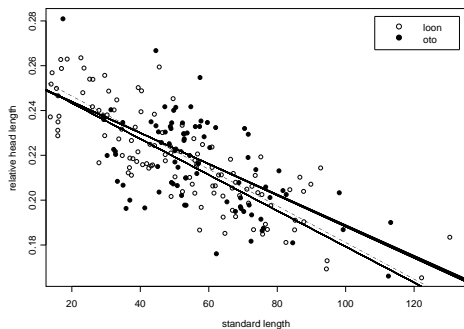

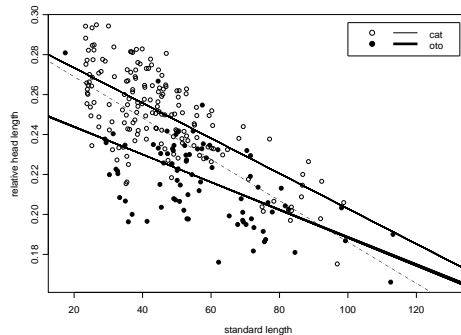Fig. 9. Deepest lines and H_0 line ($q = 2$)

Fig. 10. Deepest lines and H_0 line ($q = 2$)

If we compare the Canadian population **oto** with the European population **cat**, then we detect significant differences. The test statistic is -6.552 for $q = 2$ (and -4.733 for $q = 3$), which is in both cases less than the 0.5% quantile. Thus, the hypothesis, that both populations can be described by the same regression line has to be rejected with respect to a significance level of 1%. The deepest lines for **cat** and **oto** as well as the line, that maximizes $\Phi_\theta(z^1, z^2)$ are drawn into Figure 10.

Acknowledgement

References

[1] Dryden, I.L. and Mardia, K.V., 1999. Statistical shape analysis. John Wiley, New York.

[2] Hubert, M. and Rousseeuw, P.J., 1999. The catline for deep regression. J. Multivariate Anal. 66, 270–296.

[3] Jurečková, J. and Sen, P.K., 1996. Robust statistical procedures: asymptotics and interrelations. John Wiley, New York.

[4] Katina, S., Total variation penalty in image warping and selected multivariate methods in shape analysis. (PhD. thesis, Comenius University, Bratislava, 2004).

[5] Liu, R.Y., 1988. On a notion of simplicial depth. Proc. Nat. Acad. Sci. USA 85, 1732–1734.

[6] Liu, R.Y. , 1990. On a notion of data depth based on random simplicies. Ann. Statist. 18, 405–414.

[7] Liu, R.Y. and Singh, K., 1993. A quality index based on data depth and multivariate rank tests. J. Amer. Statist. Assoc. 88, 252–260.

[8] Marazzi, A., Object oriented S-Plus Huber type robust regression. (Technical report, 1997).

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

[10] Mosler, K., 2002. Multivariate dispersion, central regions and depth, the lift zonoid approach. Lecture Notes in Statistics, vol. 165, Springer, New York.

[11] Müller, Ch.H., 2005. Depth estimators and tests based on the likelihood principle with applications to regression. J. Multivariate Anal. 95, 153–181.

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

[13] Rousseeuw, P.J. and Hubert, M., 1999. Regression depth. J. Amer. Statist. Assoc. 94, 388–433.

[14] Rousseeuw, P.J. and Leroy, A., 1987. Robust regression and outlier detection. John Wiley, New York.

[15] Tomeček, J., Kováč, V. and Katina, S., 2005. Ontogenetic variability in external morphology of native (Canadian) and nonnative (Slovak) populations of pumpkinseed (Lepomis gibbosus, Linnaeus 1758). Journal of Applied Ichthyology. **21**, 335–344.

[16] Tukey, J.W., 1975. Mathematics and the picturing of data., In: Proceedings of the International Congress of Mathematicians 1974, vol. 2, Vancouver, 523–531.

[17] Van Aelst, S., Rousseuw, P.J., Hubert, M., and Struyf, A., 2002. The deepest regression method. J. Multivariate Anal. 81, 138–166.

[18] Wellmann, R., Harmand, P. and Müller, Ch.H., 2006. Distribution free tests for polynomial and multiple regression based on simplicial depth. Submitted.

[19] Yohai, V.J., 1987. High breakdown-point and high efficiency robust estimates for regression. Ann. Statist. 15, 642–656.