Distribution free tests for polynomial regression based on simplicial depth

by Robin Wellmann*, Peter Harmand**, Christine H. Müller*
*University of Kassel, **University of Oldenburg

July 30, 2007

Abstract

A general approach for developing distribution free tests for general linear models based on simplicial depth is presented. In most relevant cases, the test statistic is a degenerated U-statistic so that the spectral decomposition of the conditional expectation of the kernel function is needed to derive the asymptotic distribution. A general formula for this conditional expectation is derived. Then it is shown how this general formula can be specified for polynomial regression. Based on the specified form, the spectral decomposition and thus the asymptotic distribution is derived for polynomial regression of arbitrary degree. An application on cubic regression demonstrates the applicability of the new tests and in particular their outlier robustness.

Keywords: Distribution-free tests, simplicial depth, regression depth, polynomial regression, degenerated U-statistic, spectral decomposition, outlier robustness.
AMS Subject classification: Primary 62G05, 62G10; secondary 62J05, 62J12, 62G20.

1 Introduction

Simplicial depth for multivariate location was introduced by Liu (1988, 1990). It is based on the half space depth of Tukey (1975). Both depth notions lead to a generalization of the median for multivariate data which is equivariant with respect to affine transformations. Moreover the concept of depth is useful to generalize ranks.

The simplicial depth has the advantage that it is an U-statistic so that in principle the asymptotic distribution is known. However, it is not easy to derive the asymptotic distribution. Arcones et al. (1994) derived the asymptotic normality of the maximum simplicial depth estimator of Liu (1988, 1990) via the convergence of the whole U-process. The convergence of the U-process was also shown by Dümbgen (1992). However the asymptotic normal distribution has a covariance matrix which depends on the underlying distribution. Hence this result cannot be used to derive distribution-free tests. Liu (1992) and Liu and Singh (1993) proposed distribution-free multivariate rank tests which generalize the Wilcoxon's rank sum test for two samples. While the asymptotic normality is derived for several depth notions for distributions on $I\!R^1$, it is shown only for the Mahalanobis depth for distributions on $I\!R^k$, $k > 1$. Hence it is unclear how to generalize the approach of Liu and Singh to other situations.

Several other depth concepts were introduced since Tukey (1975). See for example the book of Mosler (2002) and the references in it. Multivariate depth concepts were transferred to regression by Rousseeuw and Hubert (1999), to logistic regression by Christmann and Rousseeuw (2001) and to the Michaelis-Menten model by Van Aelst et al. (2002). The depth concept for regression bases on the notion of nonfit introduced by Rousseeuw and Hubert (1999). Thereby, a regression parameter θ is called a nonfit, if there is another parameter θ' which provides for all observations z_n smaller squared residuals $r(z_n, \theta')^2$. The depth of a regression parameter θ is then given by the minimum number of observations which must be removed so that θ becomes a nonfit.

More general concepts of depth were introduced and discussed by Zuo and Serfling (2000a,b) and Mizera (2002). Mizera (2002) in particular generalized the regression depth of Rousseuw and Hubert (1999) by basing the nonfit on general quality functions instead of squared residuals. Using these quality functions, he introduced the "global depth", the "tangent depth" and the "local depth" and gave a sufficient condition for their equality. This approach makes it possible to define the depth of a parameter value with respect to given observations in various statistical models via general quality functions. Appropriate quality functions are in particular likelihood functions as studied by Mizera and Müller (2004) for a depth notion of location and scale and by Müller (2005) for a depth notion for generalized linear models.

As for multivariate location, there exist only few results concerning tests based on regression depth and its generalizations. Bai and He (1999) derived the asymptotic distribution of the maximum depth estimator for regression so that tests could be based on this. However, this asymptotic distribution is given implicitly so that it is not convenient for inference. Van Aelst et al. (2002) derived an exact test based on the regression depth of Rousseeuw and Hubert (1999) but did it only for linear regression. Müller (2005) derived tests by using the simplicial regression depth which generalizes Liu's (1988, 1990) simlicial depth to regression.

A general simplicial depth for a q dimensional parameter θ within a sample $z =$

$(z_1, ..., z_N)$ can be defined by

$$d_S(\theta, z) = \binom{N}{q+1}^{-1} \sum_{1 \le n_1 < n_2 < ... < n_{q+1} \le N} \psi_\theta(z_{n_1}, \ldots, z_{n_{q+1}}),$$

where the symmetrical kernel function $\psi_\theta \in \mathbb{L}_2(\bigotimes_{n=1}^{q+1} P_\theta^{Z_1})$ is an indicator function which equals one if the depth $d(\theta, (z_{n_1}, \ldots, z_{n_{q+1}}))$ of θ in $z_{n_1}, \ldots, z_{n_{q+1}}$ is greater than zero. If the depth d is the regression depth, the general simplicial depth is called simplicial regression depth.

Müller (2005) proposed to base the test statistic directly on the general simplicial depth d_S. For testing a hypothesis of the form $H_0 : \theta \in \Theta_0$, where Θ_0 is an arbitrary subset of the parameter space, the test statistic is defined as $T(z_1, \ldots, z_N) := \sup_{\theta \in \Theta_0} T_\theta(z_1, \ldots, z_N)$. Thereby $T_\theta(z_1, \ldots, z_N)$ is defined as

$$T_\theta(z_1, \ldots, z_N) := \frac{\sqrt{N}(d_S(\theta, (z_1, \ldots, z_N)) - \mu_\theta)}{(q+1)\ \sigma_\theta},$$

if $d_S(\theta, z)$ is not a degenerated U-statistic, i.e. $\psi_\theta^1(z_1) := E(\psi_\theta(Z_1, \ldots, Z_{q+1})|Z_1 = z_1)$ depends on z_1, and

$$T_\theta(z_1, \ldots, z_N) := N\,(d_S(\theta, (z_1, \ldots, z_N)) - \mu_\theta), \tag{1}$$

if $d_S(\theta, z)$ is a degenerated U-statistic. The null hypothesis H_0 is rejected if $T(z_1, \ldots, z_N)$ is less than the α-quantile of the asymptotic distribution of $T_\theta(Z_1, \ldots, Z_N)$. Thereby the quantities μ_θ and σ_θ are defined as $\mu_\theta = E(\psi_\theta(Z_1, \ldots, Z_{q+1}))$ and $\sigma_\theta^2 = \text{Var}(\psi_\theta^1(Z_1))$.

Unfortunately, the simplicial depth d_S is a degenerated U-statistic in the most interesting case, that the true regression function is in the center of the data, which means that the median of the residuals is zero. Whereas nondegenerated U-statistics are asymptotically normal distributed, simple asymptotic results are not possible for degenerated U-statistics. In the degenerated case, the asymptotic distributions can be derived by using the second component of the Hoeffding decomposition. We have namely the following result (see e.g. Lee 1990, p. 79, 80, 90, Witting and Müller-Funk, p. 650). If the reduced normalized kernel function

$$\psi_\theta^2(z_1, z_2) := E(\psi_\theta(Z_1, \ldots, Z_{q+1}) - \mu_\theta | Z_1 = z_1, Z_2 = z_2) \tag{2}$$

is $\mathbb{L}_2$-integrable, it has a spectral decomposition of the form

$$\psi_\theta^2(z_1, z_2) = \sum_{l=1}^{\infty} \lambda_l\ \varphi_l(z_1)\ \varphi_l(z_2), \tag{3}$$

where the functions φ_l are $\mathbb{L}_2$-integrable, normalized, and orthogonal. Then the asymptotic distribution of the simplicial depth is given by

$$N(d_S(\theta, (Z_1, \ldots, Z_N)) - \mu_\theta) \xrightarrow{\mathcal{L}} \binom{q+1}{2} \sum_{l=1}^{\infty} \lambda_l\ (U_l^2 - 1), \tag{4}$$

where $U_l \sim \mathcal{N}(0,1)$ and U_1, $U_2, \ldots$ are independent. In the general case, it could happen that the eigenvalues λ_l depend on the underlying parameter θ. However, Müller (2005) could show that this is not the case for polynomial regression in generalized linear models so that the asymptotic distribution does not depend on the regression parameter.

However, Müller (2005) was only able to find the spectral decompositions for linear and quadratic regression in generalized linear models. These spectral decompositions were found by solving differential equations. In this paper we derive the spectral decomposition (3) for polynomial regression of arbitrary degree by a complete new approach.

For this approach, we use in Section 2 a rather general quality function for defining the depth d used in the simplicial depth d_S. This general quality function, called quality function for extended linear regression, is motivated by the result of Müller (2005) that the tangent depth for quality functions based on likelihood functions in generalized linear models is based on modified residuals $g(y_n) - x(t_n)^\top\theta$, where the function g is a transformation of the dependent variables y_n and x is the regression function applied on the explanatory variables t_n. In this paper the quality function is based on $h(z_n) - v(z_n)^\top\theta$, where z_n can be $z_n = (y_n, t_n)$ with $h(z_n) = g(y_n)$ and $v(z_n) = x(t_n)$ but also other relations are possible. The simplicial depth is based on the tangent depth for this general quality function. However, this simplicial depth attain rather high values in subspaces of the parameter space, since it does not provide convex depth contours as all simplicial depth notions do not. This is in particular a disadvantage in testing if the aim is to reject the null hypothesis. To avoid this disadvantage, we introduce in Section 2 a harmonized simplicial depth for general linear models. This approach leads also to a method to calculate the maximum simplicial depth under the null hypothesis. While in Müller (2005) only null hypotheses could be rejected for which the null hypothesis is a point or a line within the parameter space, we are now able to treat hypotheses about arbitrary subspaces and polyhedrals, as Wellmann et al. (2007a) showed.

In Section 3, we derive a general formula for the conditional expectation (2) for the simplicial depth for extended linear regression models introduced in Section 2 and we show that the asymptotic distribution can be obtained by calculating the spectral decomposition of a function $\mathcal{K}$, which only depends on the probability law of the vector product of regressor variables. This means in particular that the asymptotic distribution of the test statistic (1) does not depend on the unknown regression parameter. The function $\mathcal{K}$ is applied to the harmonized form of the simplicial regression depth but the proofs hold also for the unmodified form.

The general formula for $\mathcal{K}$ is specified for polynomial regression of arbitrary degree in Section 4. Based on the specified formula, the spectral decomposition is derived. The spectral decomposition is found by a Fourier series representation of a related function of $\mathbb{L}_2[-1,1]$ which is used to derive the required representation of $\mathcal{K}$. We think that this approach can be used to find the spectral decomposition of other simplicial depth functions. In particular, Wellmann and Müller (2007b) derived the asymptotic distribution of the simplicial regression depth for different models of multiple regression.

Section 5 provides an application on tests in a cubic regression model. This example in particular shows that the new tests possess high outlier robustness. All proofs are given in Section 6.

2 Simplicial depth for extended linear regression

We assume that the random vectors $Z_1, ..., Z_N$ are independent and identically distributed throughout the paper. The random vectors Z_n have values in $\mathcal{Z} \subset I\!R^p$, $p \geq 1$. There exist known functions $v : \mathcal{Z} \to I\!R^q$ and $h : \mathcal{Z} \to I\!R$ so that the random variables $X_n = v(Z_n)$ and $Y_n = h(Z_n)$ satisfy

$$Y_n = X_n^T \theta + E_n$$

for $\theta \in \Theta = I\!R^q$ and random error E_n. Random variables are denoted by capital letters and realizations by small letters. The value $s_n(\theta) := \text{sign}_\theta(z_n) := \text{sign}(y_n - x_n^T \theta)$ is called the sign of the residual of the n-th (transformed) observation. The family $\mathcal{P} = \{P_\theta^{(Z_1,...,Z_N)} : \theta \in \Theta\}$ of probability measures with $\Theta = I\!R^q$ may be unknown, but for the purpose of deriving tests, we will assume that the following assumptions hold:

- $P_\theta(S_1(\theta) = 1|X_1) \equiv \frac{1}{2}$ a.s., (5)
- $P_\theta(S_1(\theta) = 0|X_1) \equiv 0$ a.s., and
- $P_\theta(X_1, \ldots, X_q$ are linearly dependent$) = 0$.

While the last two conditions of (5) are usually satisfied for continuous distributions, the first condition can be satisfied by appropriate transformations v and h. In a model satisfying (5), the following quality function can be used.

Definition 1 (Quality functions for extended linear regression) *For $z_n \in \mathcal{Z}$ take φ_{z_n} to be a function with continuous derivatives, which has its maximum and its sole critical point in* 0. *Let $v : \mathcal{Z} \to I\!R^q$ and $h : \mathcal{Z} \to I\!R$ be measurable.*
Then the function

$$\mathcal{G}_{z_n} : \Theta \to I\!R \text{ with } \mathcal{G}_{z_n}(\theta) := \varphi_{z_n}(h(z_n) - v(z_n)^T \theta)$$

is said to be a quality function for extended linear regression.

Mostly, one would choose the likelihood functions to be the quality functions. However, for the resulting transformations, not always $p := P_\theta(E_n > 0|X_n) = \frac{1}{2}$ is satisfied, so that the true regression function is not in the center of the transformed data. In these cases it is more appropriate to choose the transformation h and v, such that $p = \frac{1}{2}$ holds and to define the quality functions by

$$\mathcal{G}_{z_n}(\theta) := -(h(z_n) - v(z_n)^T\theta)^2.$$

Although quality functions are needed to define the tangent depth or the global depth of Mizera (2002), the resulting depth functions do not depend on the choice of φ_{z_n}, so that we may restrict ourselves to the simplest case $\varphi_{z_n}(x) = -x^2$.

The following example demonstrates, how h and v and thus the quality function can be obtained: For generalized linear models the observations are given by $Z_n = (U_n, T_n)$ which have density $f_\theta^{(U_n,T_n)}(u_n, t_n) = f_\theta^{U_n|x(T_n)=x(t_n)}(u_n) f^{T_n}(t_n)$. Although these likelihood functions are quality functions for extended linear regression with regressors $v(z_n) = x(t_n)$ in most models for general linear regression, not always the assumption $p = \frac{1}{2}$ is satisfied. For example, for regression with exponential distributed dependent observations, that is,

$$f_\theta^{U_n|x(T_n)=x(t_n)}(u_n) = \lambda_n \exp(-\lambda_n u_n) \text{ with } \lambda_n = \exp\big(-x(t_n)^T\theta\big),$$

the approach via likelihood functions leads to $Y_n = h(Z_n) = log(U_n)$, whereas only the transformation $Y_n = h(Z_n) = log(\frac{U_n}{log(2)})$ leads to $p = \frac{1}{2}$.

Definition 2 (Tangent depth) *According to Mizera (2002), we define the tangent depth of $\theta \in \Theta$ with respect to given observations $z_1, ..., z_N \in \mathcal{Z}$ to be*

$$d_T(\theta, z) = \min_{u \neq 0} \#\{n : u^T \nabla \mathcal{G}_{z_n}(\theta) \geq 0\},$$

where $\mathcal{G}_{z_1}, ..., \mathcal{G}_{z_N}$ are quality functions for extended linear regression, $z := (z_1, ..., z_N)$ and $\nabla \mathcal{G}_{z_n}(\theta)$ denotes the vector of partial derivatives of $\mathcal{G}_{z_n}$ in θ.

As shown in Mizera (2002), this depth notion counts the number of observations that needs to be removed such that there is a "better" parameter for all remaining observations. It's easy to see, that the tangent depth does not depend on the choice of φ_{z_n}. Furthermore, for all $\theta \in \Theta$ and for given observations $z_1, ..., z_N \in \mathcal{Z}$ we have:

$$d_T(\theta, z) = \min_{u \neq 0} \#\{n : s_n(\theta)\, u^T x_n \geq 0\}. \tag{6}$$

As in Rousseuw and Hubert (1999) it can be shown, that the parameter space $\Theta = \mathbb{R}^q$ is divided up into domains with constant depth by the hyperplanes

$$H_n = \{\theta \in \mathbb{R}^q : s_n(\theta) = 0\}, n = 1, ..., N.$$

For given observations let $\text{Dom}(z)$ be the set of all those domains. We define $\bar{d}_T(G, z) := d_T(\theta, z)$ for $G \in \text{Dom}(z)$ and $\theta \in G$.

We will define the simplicial depth to be an U-statistic. If we would take the tangent depth to be the kernel function of the U-statistic, then the simplicial regression depth attain rather high values in subspaces of the parameter space, namely in $\text{Border}(z) := \cup_{n=1}^{N} H_n$. This is in particular a disadvantage if the aim is to reject the null hypothesis. To avoid this disadvantage, we introduce a harmonized depth.

Definition 3 (Harmonized depth) *The harmonized depth of $\theta \in \Theta$ with respect to the observations $z_1, ..., z_N \in \mathcal{Z}$ is defined to be*

$$\psi_\theta(z) = \min_{G \in \text{Dom}(z), \theta \in \bar{G}} \bar{d}_T(G, z),$$

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

Definition 4 (Simplicial depth) *The simplicial depth is given by*

$$d_S(\theta, z) = \binom{N}{q+1}^{-1} \sum_{1 \le n_1 < n_2 < \ldots < n_{q+1} \le N} \psi_\theta(z_{n_1}, \ldots, z_{n_{q+1}}).$$

This depth, which transfers the simplicial depth of Liu to regression models, is also called a simplicial depth because it counts the fraction of simplicies that are bounded by $q+1$ hyperplanes from $H_1, ..., H_N$ and contain θ as an interiour point. Algorithms for calculating the simplicial depth are based on this view as well and are given in Wellmann et al. (2007a). The proposed tests are based on the asymptotic distribution of this depth notion.

3 The asymptotic distribution of the simplicial depth

The definition of tangent depth shows, that the depth of a parameter is the halfspace depth of 0 with respect to the gradients of the quality functions. Thereby, the half space depth of 0 with respect to given vectors $r_1, \ldots, r_N \in I\!R^q$ is defined as

$$d_H(0, r) := \min_{u \neq 0} \#\{n : u^T r_n \ge 0\},$$

where $r = (r_1, ..., r_{q+1})$ (see Tukey, 1975). The next lemma is needed to derive the conditional expectations of the kernel function, which depend only on $q+1$ observations, but it can also be used to calculate the simplicial depth of a given parameter.

Lemma 1 *Let $r_1, \ldots, r_{q+1} \in I\!R^q$ be in general position. Then $d_H(0,r) \in \{0,1\}$ and the following statements are equivalent:*

(i) $d_H(0,r) = 0$,

(ii) $r_1 \notin I\!R_{\leq 0} r_2 + \ldots + I\!R_{\leq 0} r_{q+1}$.

Proofs are given in the appendix. The next Proposition shows, that

$$\psi_\theta^1(z_1) := E(\psi_\theta(Z_1, \ldots, Z_{q+1})|Z_1 = z_1)$$

does not depend on z_1, so that the simplicial depth is a degenerated U-statistic and has asymptotically the distribution of an infinite linear combination of χ^2-distributed random variables (see e.g. Lee 1990, p. 79, 80, 90, Witting and Müller-Funk, p. 650). This distribution depends only on the conditional expectation

$$\psi_\theta^2(z_1, z_2) := E(\psi_\theta(Z_1, \ldots, Z_{q+1})|Z_1 = z_1, Z_2 = z_2) - E(\psi_\theta(Z_1, \ldots, Z_{q+1})).$$

Proposition 1 *Let $\theta \in \Theta$ and let $z_1, z_2 \in \mathcal{Z}$, such that x_1, x_2 are linearly independent and $s_1(\theta), s_2(\theta) \in \{-1, 1\}$. Then*

$$\psi_\theta^1(z_1) = \frac{1}{2^q}$$

and

$$\psi_\theta^2(z_1, z_2) = \frac{s_1(\theta) s_2(\theta)}{2^{q-1}} \left(P_\theta(x_1^T W \, x_2^T W < 0) - \frac{1}{2} \right),$$

where $W := X_3 \times \ldots \times X_{q+1}$ is the vector product of $X_3, \ldots, X_{q+1}$.

With this proposition, we obtain a main result: We get the asymptotic distribution of the simplicial depth in extended linear regression by calculating the spectral decomposition of the kernel $\mathcal{K}$, defined by

$$\mathcal{K}(x_1, x_2) := P_\theta(x_1^T W \, x_2^T W < 0) - \frac{1}{2}, \quad \text{for } x_1, x_2 \in I\!R^q. \tag{7}$$

Note that $x_i^\top W = \det(x_i, X_3, \ldots, X_{q+1})$ for $i = 1, 2$. The spectral decomposition is a representation

$$\mathcal{K}(x_1, x_2) = \sum_{j=1}^{\infty} \lambda_j \varphi_j(x_1) \varphi_j(x_2) \quad \text{in } I\!L_2\big(P^{X_1} \otimes P^{X_1}\big),$$

where $(\varphi_j)_{j=1}^\infty$ is an orthonormal system (ONS) in $I\!L_2\big(P^{X_1}\big)$ and $\lambda_1, \lambda_2, \ldots \in I\!R$. The functions $(\varphi_j)_{j=1}^\infty$ are eigenfunctions and the values $\lambda_1, \lambda_2, \ldots$ are the corresponding eigenvalues of the related integral operator $T_\mathcal{K}$, defined by

$$T_\mathcal{K} : I\!L_2(P^{X_1}) \to I\!L_2(P^{X_1}) \text{ with } T_\mathcal{K} f(s) = \int \mathcal{K}(s,t) f(t) \, dP^{X_1}(t).$$

The system $(\psi_j)_{j=1}^{\infty}$, defined by $\psi_j(z) := \mathrm{sign}_\theta(z)\ \varphi_j(v(z))$ for $z \in \mathcal{Z}$ is an ONS in $I\!L_2(P^{Z_1})$ and for ψ_θ^2 we have the representation

$$\begin{aligned}\psi_\theta^2(z_1, z_2) &= \frac{\mathrm{sign}_\theta(z_1)\mathrm{sign}_\theta(z_2)}{2^{q-1}}\,\mathcal{K}(v(z_1), v(z_2))\\ &= \sum_{j=1}^{\infty} \frac{1}{2^{q-1}}\lambda_j\ \mathrm{sign}_\theta(z_1)\varphi_j(v(z_1))\ \mathrm{sign}_\theta(z_2)\varphi_j(v(z_2))\\ &= \sum_{j=1}^{\infty} \frac{1}{2^{q-1}}\lambda_j\ \psi_j(z_1)\ \psi_j(z_2).\end{aligned}$$

Hence, it follows by (4), that

$$N\big(d_S(\theta, (Z_1, \dots, Z_N)) - \frac{1}{2^q}\big) \xrightarrow{\mathcal{L}} \sum_{l=1}^{\infty} \frac{(q+1)!}{(q-1)!2^q}\lambda_l\ \big({U_l}^2 - 1\big), \tag{8}$$

where $U_1, U_2, \dots$ are i.i.d. random variables with $U_1 \sim \mathcal{N}(0,1)$. Furthermore, this derivation shows, that the asymptotic distribution does not depend on the underlying parameter θ, if the distribution of W does not depend on it. In the next section, this general result is applied to polynomial regression.

4 Polynomial regression

A special extended linear regression model is the polynomial regression model of degree $r = q-1$. In this model, the unknown parameter is $\theta = (\theta_1, \dots, \theta_q)^\top \in I\!R^q$, $\mathcal{Z} \subset I\!R^2$, $Z_n = (Y_n, T_n)$, the vector $v(Z_n) := x(T_n) := (1, T_n, ..., T_n^r)^T$ is the regressor and $h(Z_n) = Y_n$ is the dependent variable. Because of the independence of $T_1, ..., T_N$, the third assumption in (5) is equivalent to $P_\theta(T_1 = t) = 0$ for all $t \in I\!R$.

In this section, we derive the asymptotic distribution of the simplicial depth by calculating the spectral decomposition of the kernel $\mathcal{K}$, given in (7). While Müller (2005) derived it only for $r = 1$ and $r = 2$ in another way, we have now the asymptotic distribution for polynomial regression of arbitrary degree. At first, we give a simple representation of the kernel $\mathcal{K}$, whis is obtained from (7) via the formula for Vandermonde determinants (see the appendix).

Proposition 2 *For all $t_1, t_2 \in I\!R$ we have*

$$\mathcal{K}(x(t_1), x(t_2)) = -2^{r-1}\left(\frac{1}{2} - |F^{T_1}(t_1) - F^{T_1}(t_2)|\right)^r,$$

where F^{T_1} is the distribution function of T_1.

Müller derived the same formula for the reduced normalized kernel function ψ_θ^2 (see Proposition 2 in Müller 2005). Our proof is based not on ψ_θ^2 , but on $\mathcal{K}$. This makes the proof much shorter. It remains to derive the spectral decomposition of $\mathcal{K}$, which we obtain in the next proposition via a Fourier series representation of $(\frac{1}{2} - |z|)^r$ in $\mathbb{L}_2[-1, 1]$.

Proposition 3 *The spectral decomposition of* $\left(\frac{1}{2} - |s-t|\right)^r$ *in* $\mathbb{L}_2[0,1]^2$ *is given by*

$$\left(\frac{1}{2} - |s-t|\right)^r = \gamma_0^{(r)} \cdot 1 + \sum_{l=1}^{\infty} \gamma_l^{(r)} \cdot 2 \cdot [\cos(k\pi s)\ \cos(k\pi t) + \sin(k\pi s)\ \sin(k\pi t)]$$

where for r *odd*

$$\gamma_l^{(r)} = \begin{cases} 0, & \text{if } l \text{ is even,} \\ -\sum\limits_{\substack{k \in \{1,\dots,r\} \\ k \text{ odd}}} \frac{r!}{2^{r-k-1}(r-k)!} (-l^2 \pi^2)^{-\frac{k+1}{2}}, & \text{if } l \text{ is odd,} \end{cases}$$

and for r *even*

$$\gamma_l^{(r)} = \begin{cases} \frac{1}{(r+1)\, 2^r}, & \text{if } l = 0, \\ -\sum\limits_{\substack{k \in \{1,\dots,r\} \\ k \text{ odd}}} \frac{r!}{2^{r-k-1}(r-k)!} (-l^2 \pi^2)^{-\frac{k+1}{2}}, & \text{if } l \text{ is even and } l > 0, \\ 0, & \text{if } l \text{ is odd.} \end{cases}$$

Let $(\psi_j)_{j\in J}$ be the ONS in $\mathbb{L}_2[0, 1]$, given in the proof of Proposition 3, such that $(\gamma_j^{(r)})_{j\in J}$ are the eigenvalues, related to $K(s, t) = (\frac{1}{2} - |s - t|)^r$. Then the system $(\varphi_j)_{j\in J}$, defined by $\varphi_j := \psi_j \circ F^{T_1} \circ x^{-1}$ is an ONS in $\mathbb{L}(P^{X_1})$ and we have the representation

$$\begin{aligned} \mathcal{K}(x_1, x_2) &= \mathcal{K}(x(x^{-1}(x_1)), x(x^{-1}(x_2))) \\ &= -2^{r-1}(\frac{1}{2} - |F^{T_1}(x^{-1}(x_1)) - F^{T_1}(x^{-1}(x_2))|)^r \\ &= \sum_{j\in J} (-2^{r-1}\gamma_j^{(r)})\varphi_j(x_1)\varphi_j(x_2). \end{aligned}$$

Hence, the next Theorem holds:

Theorem 1 *If $P(Y_n - x(T_n)^\top \theta \geq 0 | T_n) = \frac{1}{2}$ and T_n has a continuous distribution, then the asymptotic distribution of the simplicial likelihood depth $d_S(\theta, (Z_1, \ldots, Z_N))$ for polynomial regression is given by*

$$N\left(d_S(\theta, (Z_1, \ldots, Z_N)) - \frac{1}{2^{r+1}}\right) \xrightarrow{\mathcal{L}} \sum_{l=0}^{\infty} \lambda_{2l+1}(V_l^2 + W_l^2 - 2)$$

for r even and

$$N\left(d_S(\theta, (Z_1, \ldots, Z_N)) - \frac{1}{2^{r+1}}\right) \xrightarrow{\mathcal{L}} \lambda_0(U^2 - 1) + \sum_{l=1}^{\infty} \lambda_{2l}(V_l^2 + W_l^2 - 2)$$

for r odd, where U, $V_0, W_0, V_1, W_1, \ldots$ are independent random variables with standard normal distribution and

$$\begin{aligned} \lambda_0 &= -\frac{r+2}{2^{r+2}}, \\ \lambda_l &= \sum_{\substack{k \in \{1,\ldots,r\} \\ k \text{ odd}}} \frac{(r+2)!}{2^{r-k+1}(r-k)!}(-l^2\pi^2)^{-\frac{k+1}{2}} \quad \text{for } l \in I\!N. \end{aligned}$$

The calculation of the test statistic and the critical values for any hypothesis of the form $H_0 : \theta \in \Theta_0$ where Θ_0 is a subspace of the parameter space or a polyhedron is described in Wellmann et al. (2007a). There also a table of the critical values is given. The example in Section 5 shows the applicability of the method for $q > 3$, whereas in Müller (2005) and Wellmann et al. (2007a) examples for $q \leq 3$ were given. In particular, in Wellmann et al. an example is calculated, where the hypotheses is given by a polyhedral. There it is also shown that the method can be used for tests in two sample problems. The examples demonstrate that the tests are outlier robust. See also Wellmann (2007c).

5 Application: Test about quadratic function against cubic function

The concentration of malondialdehyd (MDA) for 78 women twice after childbirth (IMDA and IIMDA) at two time points was measured, to find a relation between the levels of IMDA and IIMDA. MDA is a metabolite of lipid peroxides detectable in plasma. It was measured as an indicator of lipid peroxidation and oxidation stress of women post partum (after childbirth). The data came from the Clinic of Gynaecology, Faculty Hospital with Policlinic, Bratislava-Ružinov (Slovakia).

We assume a cubic regression model ($q = 4$) and choose t_n as IIMDA and y_n as IMDA. Normality of the residuals with respect to the ordinary least square estimation

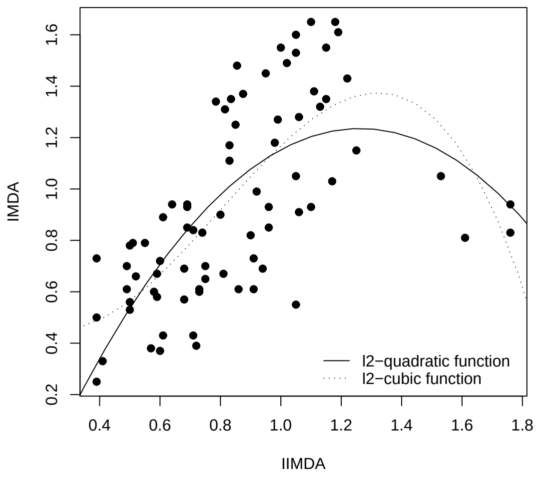

Figure 1: Least squares quadratic and cubic function

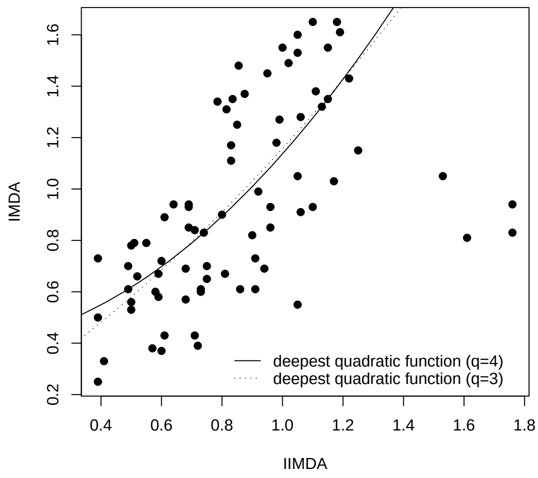

Figure 2: Deepest quadratic functions for $q = 3$ and $q = 4$.

was rejected (p-value< 0.001) by the χ^2 goodness-of-fit test (function `chisq.gof` in S-Plus). Assuming a cubic regression model with parameter $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)^T \in I\!R^q$, we want to test the hypothesis that the true function is quadratic. Hence, we want to test $H_0 : \theta \in \Theta_0$ with $\Theta_0 = \{\theta \in I\!R^4 : \theta_4 = 0\}$. In the cubic regression model with $q = 4$ and $\theta_4 = 0$, a parameter with maximum simplicial depth is $\widehat{\theta}_D = (0.3942166, 0.150213, 0.592007, 0)^T$ (Figure 2).

The maximal simplicial depth is $\sup_{\theta\in\Theta_0} d_S(\theta, (z_1, ..., z_{78})) = 0.067$ and the test statistic is $T(z_1, ..., z_{78}) = 0.364$, which is more than the 90%-quantile (see Wellmann et al. 2007a) and hence we can not reject the null hypothesis for the significance level 10%. Thus, we may assume a model for quadratic regression ($q = 3$). The deepest quadratic function in the quadratic regression model is rather similar to the deepest quadratic function in the cubic regression model (see Figure 2). Note, that also the hypothesis, that the true function is a linear function, cannot be rejected within the model for quadratic regression ($q = 3$), since the test statistic is 0.534, which is also more than the 90%-quantile of the asymptotic distribution.

The least squares estimation within the model for quadratic regression is $\widehat{\theta}_{l_2} = (-0.684688, 3.043423, -1.206010)^T$. Within the model for cubic regression it is $\widehat{\theta}_{l_2} = (0.676493, -1.739798, 3.834566, -1.602871)^T$ (Figure 1). Although the residuals are not normally distributed, we tests the null hypothesis $H_0 : \theta_4 = 0$ against $H_1 : \theta_4 \neq 0$ by the F-test (function `anova` in S-Plus). The F-test provides for the null hypothesis a p-value=0.018, so $H_0 : \theta_4 = 0$ is rejected. This is due to the outliers at the right hand side and without them a quadratic or linear regression function is a good description of the data.

6 Proofs

For more details of the proofs see also Wellmann (2007c).

Proof of Lemma 1

Since $r_1, ..., r_q$ are linearly independent, they belong to a hyperplane H with $0 \notin H$. There is a $\gamma < 0$ and a $u \in I\!R^q$, such that $H = \{v \in I\!R^q : v^T u = \gamma\}$.
Since $r_1, ..., r_q$ don't belong to the half space $\{v \in I\!R^q : v^T u \geq 0\}$, we have $d_H(0, r) \leq 1$. It remains to show the equivalence.

(ii)$\Rightarrow$ (i): For any $j = 1, \ldots, q+1$ let H_j be the hyperplane that contains the points $(r_i)_{i \in \{1,\ldots,q+1\} \setminus \{j\}}$.

Step 1: There is a $j \in \{1, \ldots, q+1\}$ such that 0 and r_j are on different sides of H_j.

Proof: Since $r_2, \ldots, r_{q+1}$ is a Basis of $I\!R^q$, there exists $\gamma_2, \ldots, \gamma_{q+1} \in I\!R$ such that $r_1 = \gamma_2 r_2 + ... + \gamma_{q+1} r_{q+1}$. Since $r_1 \notin I\!R_{\leq 0} r_2 + ... + I\!R_{\leq 0} r_{q+1}$ we may assume that $\gamma_2 > 0$. We prove that r_1 and 0 are on different sides of H_1, if r_2 and 0 are on the same side of H_2. Hence, we have:

- **(a)** $r_1 = \gamma_2 r_2 + ... + \gamma_{q+1} r_{q+1}$ with $\gamma_2 > 0$.
- **(b)** There are $\alpha > 0$ and $\beta_3, \ldots, \beta_{q+1} \in I\!R$ such that $r_2 = r_1 - \alpha\, r_1 + \sum_{j=3}^{q+1} \beta_j (r_j - r_1) \in I\!R_{<0} r_1 + H_2$.

From these equations we obtain two different representations of r_2:

$$\begin{aligned} r_2 &= \Big(1 - \alpha - \sum_{j=3}^{q+1} \beta_j\Big) r_1 + \beta_3 r_3 + ... + \beta_{q+1} r_{q+1}, \\ r_2 &= \frac{1}{\gamma_2} r_1 + \frac{-\gamma_3}{\gamma_2} r_3 + ... + \frac{-\gamma_{q+1}}{\gamma_2} r_{q+1}. \end{aligned}$$

Comparing the coefficients leads to

$$0 < \frac{1}{\gamma_2} = 1 - \alpha - \sum_{j=3}^{q+1} \beta_j \;\text{ and }\; \frac{-\gamma_k}{\gamma_2} = \beta_k \;\text{ for }\; k = 3, \ldots, q+1.$$

It follows that

$$\begin{aligned}\gamma_2 + ... + \gamma_{q+1} &= \gamma_2 - \gamma_2\beta_j - ... - \gamma_2\beta_{q+1} \\ &= \gamma_2\big(1 - \sum_{j=3}^{q+1}\beta_j\big) \\ &= \frac{1 - \sum_{j=3}^{q+1}\beta_j}{1 - \alpha - \sum_{j=3}^{q+1}\beta_j} > 1.\end{aligned}$$

With (a) we have

$$r_1 = (\gamma_2 + \gamma_3 + ... + \gamma_{q+1})r_2 + \gamma_3(r_3 - r_2)... + \gamma_{q+1}(r_{q+1} - r_2).$$

Thus there is a $\lambda \in (0,1)$ with: $\lambda\, r_1 \in r_2 + \sum_{j=3}^{q+1} I\!R(r_j - r_2) = H_1$.
Hence, r_1 and 0 are on different sides of H_1. This finishes the proof of Step 1.

Step 2: Main proof. The vectors r_j and 0 are on different sides of this affine hyperplane H_j. Let $v \in H_j$. All vectors $r_1, \ldots, r_{q+1}$ are in the open half space $I\!R_{>0}v + (H_j - v)$. The half space $I\!R^q \backslash (I\!R_{>0}v + (H_j - v))$ don't contain the vectors $r_1, \ldots, r_{q+1}$. Thus, $d_H(0, r) = 0$.

(i)⇒(ii)

The vectors $r_1, \ldots, r_{q+1}$ belong to an open half space H with $0 \in \partial\, H$

$$\begin{aligned}&\Rightarrow\ -r_2, \ldots, -r_{q+1} \in H' := I\!R^q \backslash H \\ &\Rightarrow\ I\!R_{\leq 0}r_2, \ldots, I\!R_{\leq 0}r_{q+1} \subset H' \\ &\Rightarrow\ I\!R_{\leq 0}r_2 + ... + I\!R_{\leq 0}r_{q+1} \subset H'.\end{aligned}$$

Because of $r_1 \notin H'$ it follows that $r_1 \notin I\!R_{\leq 0}r_2 + ... + I\!R_{\leq 0}r_{q+1}$. □

Proof of Proposition 1
Let $z_1, ..., z_m \in \mathcal{Z}$, such that $x_1, ..., x_m$ are linearly independent and $s_n := s_n(\theta) \neq 0$ for $n = 1, ..., m$, where $m \leq q$. Let be

$$\begin{aligned}\tilde{X} &:= (X_{m+1}, ..., X_{q+1}), \\ Z' &:= (z_1, ..., z_m, Z_{m+1}, ..., Z_{q+1}), \\ \mathcal{X} &:= v(\mathcal{Z}), \\ \boldsymbol{X}_{gp} &:= \{(x_{m+1}, ..., x_{q+1}) \in \mathcal{X}^{q+1-m} : \text{each subset of } q \text{ vectors from } x_1, ..., x_{q+1} \text{is lin. indep.}\}, \\ \boldsymbol{X}_{ex} &:= \{(x_{m+1}, ..., x_{q+1}) \in \boldsymbol{X}_{gp} : \exists s_{m+1}, ..., s_{q+1} \in \{-1, 1\} : d_H(0, (s_1x_1, ..., s_{q+1}x_{q+1})) = 1\}.\end{aligned}$$

We have to calculate $E(\psi_\theta(Z_1,...,Z_{q+1})|Z_1=z_1,...,Z_m=z_m)=E(\psi_\theta(Z'))$.
Since $P(\tilde{X}\in \boldsymbol{X}_{gp})=1$ and $P(\psi_\theta(Z')\in\{0,1\})=1$, we have

$$E(\psi_\theta(Z'))=P(\psi_\theta(Z')=1 \text{ and } \tilde{X}\in \boldsymbol{X}_{gp}).$$

Because of $\{\psi_\theta(Z')=1\}\cap\{\tilde{X}\in\boldsymbol{X}_{gp}\}\subset\{\tilde{X}\in\boldsymbol{X}_{ex}\}$,
it follows that

$$E(\psi_\theta(Z'))=P(\psi_\theta(Z')=1 \text{ and } \tilde{X}\in \boldsymbol{X}_{ex}).$$

For $r_2,...,r_{q+1}\in I\!R^q$ let $\sigma_{(r_2,...,r_{q+1})}:\{2,...,q+1\}\to\{-1,1\}$, such that

$$s_1x_1\in I\!R_{\le 0}\sigma_{(r_2,...,r_{q+1})}(2)r_2+...+I\!R_{\le 0}\sigma_{(r_2,...,r_{q+1})}(q+1)r_{q+1},$$

if each subset of q vectors from $s_1x_1,r_2,...,r_{q+1}$ is linearly independent.
Since $x_2,...,x_m$ are fixed, we can write $\sigma_{\tilde{x}}:=\sigma_{(x_2,...,x_{q+1})}$ for $\tilde{x}=(x_{m+1},...,x_{q+1})\in\boldsymbol{X}_{gp}$.

Now, we prove that

$$P(\psi_\theta(Z')=1 \text{ and } \tilde{X}\in\boldsymbol{X}_{ex})=P(\forall n=m+1,...,q+1:\text{sign}_\theta(Z_n)=\sigma_{\tilde{X}}(n),\tilde{X}\in\boldsymbol{X}_{ex}\}.$$

Therefore let $z_{m+1},...,z_{q+1}\in\mathcal{Z}$ with $\tilde{x}:=(x_{m+1},...,x_{q+1})\in\boldsymbol{X}_{ex}$.
Since $\tilde{x}\in\boldsymbol{X}_{ex}$ there are $s_{m+1},...,s_{q+1}\in\{-1,1\}$ with $d_H(0,(s_1x_1,...,s_{q+1}x_{q+1}))=1$.
With Lemma 1 it follows that $s_1x_1\in I\!R_{\le 0}s_2x_2+...+I\!R_{\le 0}s_{q+1}x_{q+1}$.
Hence, the definition of $\sigma_{\tilde{x}}$ implies that

$$s_n=\sigma_{\tilde{x}}(n) \text{ for } n=2,...,m. \tag{9}$$

Furthermore, we have

$$\begin{array}{cl} & \psi_\theta(z)=1\\ \Leftrightarrow & \theta\notin\cap_{n=1}^{q+1}H_n \text{ and } d_T(\theta,z)=1\\ \Leftrightarrow & \text{sign}_\theta(z_n)\neq 0 \text{ for } n=1,...,q+1 \text{ and } d_T(\theta,z)=1\\ \Leftrightarrow & \text{sign}_\theta(z_n)\neq 0 \text{ for } n=1,...,q+1 \text{ and } d_H(0,(s_1(\theta)x_1,...,s_{q+1}(\theta)x_{q+1}))=1\\ \overset{\text{Lemma 1}}{\Leftrightarrow} & \text{sign}_\theta(z_n)\neq 0 \text{ for } n=1,...,q+1 \text{ and}\\ & \text{sign}_\theta(z_1)x_1\in I\!R_{\le 0}\text{sign}_\theta(z_2)x_2+...+I\!R_{\le 0}\text{sign}_\theta(z_{q+1})x_{q+1}\\ \Leftrightarrow & \forall n=2,...,q+1:\text{sign}_\theta(z_n)=\sigma_{\tilde{x}}(n)\\ \overset{(9)}{\Leftrightarrow} & \forall n=m+1,...,q+1:\text{sign}_\theta(z_n)=\sigma_{\tilde{x}}(n). \end{array}$$

Hence,

$$\begin{aligned} E(\psi_\theta(Z')) &= P(\forall n=m+1,...,q+1:\text{sign}_\theta(Z_n)=\sigma_{\tilde{X}}(n),\tilde{X}\in\boldsymbol{X}_{ex})\\ &= \int_{\boldsymbol{X}_{ex}} P(\forall n=m+1,...,q+1:\text{sign}_\theta(Z_n)=\sigma_{\tilde{x}}(n)|\tilde{X}=\tilde{x})\,dP^{\tilde{X}}(\tilde{x}). \end{aligned}$$

Since $(\text{sign}_\theta(Z_{m+1}), ..., \text{sign}_\theta(Z_{q+1}))$ and $\tilde{X}$ are independent, it follows, that

$$\begin{aligned} E(\psi_\theta(Z')) &= \int_{\boldsymbol{X}_{ex}} P(\forall n = m+1, ..., q+1 : \text{sign}_\theta(Z_n) = \sigma_{\tilde{x}}(n))\, d\, P^{\tilde{X}}(\tilde{x}) \\ &= \int_{\boldsymbol{X}_{ex}} \prod_{n=m+1}^{q+1} P(\text{sign}_\theta(Z_n) = \sigma_{\tilde{x}}(n))\, d\, P^{\tilde{X}}(\tilde{x}) \\ &= \int_{\boldsymbol{X}_{ex}} \left(\frac{1}{2}\right)^{q+1-m} d\, P^{\tilde{X}}(\tilde{x}) \\ &= \left(\frac{1}{2}\right)^{q+1-m} P(\tilde{X} \in \boldsymbol{X}_{ex}). \end{aligned}$$

For $m = 1$ we have $\boldsymbol{X}_{gp} \subset \boldsymbol{X}_{ex}$ and thus $\psi_\theta^1(z_1) = \left(\frac{1}{2}\right)^{q+1-1} P(\tilde{X} \in \boldsymbol{X}_{ex}) = \frac{1}{2^q}$.
It remains to prove the second equation. Therefore, let $m = 2$.

Let $x_3, ..., x_{q+1} \in \mathcal{X}$, such that $(x_3, ..., x_{q+1}) \in \boldsymbol{X}_{gp}$ and let $w := x_3 \times ... \times x_{q+1}$. Then we have

$$\begin{aligned} & (x_1, ..., x_{q+1}) \in \boldsymbol{X}_{ex} \\ \overset{\text{Def.}}{\Leftrightarrow} \quad & \exists s_3, ..., s_{q+1} \in \{-1, 1\} : d_H(0, (s_1 x_1, ..., s_{q+1} x_{q+1})) = 1 \\ \overset{\text{Prop.1}}{\Leftrightarrow} \quad & \exists s_3, ..., s_{q+1} \in \{-1, 1\} : s_1 x_1 \in \mathbb{R}_{<0} s_2 x_2 + ... + \mathbb{R}_{<0} s_{q+1} x_{q+1} \\ \Leftrightarrow \quad & \exists \alpha, \beta > 0, \exists \lambda \in \mathbb{R}^q, \lambda \neq 0 : (\alpha s_1 x_1 + \beta s_2 x_2, x_3, ..., x_{q+1})\lambda = 0 \\ \Leftrightarrow \quad & \exists \alpha, \beta > 0 : \det(\alpha s_1 x_1 + \beta s_2 x_2, x_3, ..., x_{q+1}) = 0 \\ \Leftrightarrow \quad & \exists \alpha, \beta > 0 : (\alpha s_1 x_1 + \beta s_2 x_2)^T w = 0 \\ \Leftrightarrow \quad & \exists \alpha, \beta > 0 : \alpha s_1 x_1^T w + \beta s_2 x_2^T w = 0 \\ \Leftrightarrow \quad & \text{sign}(s_1 x_1^T w) = -\text{sign}(s_2 x_2^T w) \\ \Leftrightarrow \quad & s_1 s_2 x_1^T w x_2^T w < 0. \end{aligned}$$

Note, that the equation

$$P(sU < 0) = sP(U < 0) + \frac{1-s}{2}$$

holds for each $\mathbb{R}$-valued random variable U with $P(U = 0) = 0$ and $s \in \{-1, 1\}$. It follows that

$$
\begin{aligned}
\psi_\theta^2(z_1, z_2) &= E(\psi_\theta(Z')) - E(\psi_\theta) \\
&= \left(\frac{1}{2}\right)^{q+1-2} P(\tilde{X} \in \boldsymbol{X}_{ex}) - \frac{1}{2^q} \\
&= \left(\frac{1}{2}\right)^{q-1} P(s_1 s_2 x_1^T W x_2^T W < 0) - \frac{1}{2^q} \\
&= \left(\frac{1}{2}\right)^{q-1} (s_1 s_2 P(x_1^T W x_2^T W < 0) + \frac{1 - s_1 s_2}{2}) - \frac{1}{2^q} \\
&= \frac{s_1 s_2 (P(x_1^T W x_2^T W < 0) - \frac{1}{2})}{2^{q-1}}.
\end{aligned}
$$

Proof of Proposition 2
Note, that the equation

$$
P\Big(\prod_{j=1}^{N} U_j < 0\Big) = \frac{1}{2} - \frac{1}{2}(1 - 2P(U_1 < 0))^N
$$

holds for $N \in I\!N$ and i.i.d. $I\!R$-valued random variables $U_1, ..., U_N$ with $P(U_1 = 0) = 0$. Since the occurring determinants are Vandermonde determinants, we have for all $t_1, t_2 \in I\!R$:

$$
\begin{aligned}
&\mathcal{K}(x(t_1), x(t_2)) \\
&\quad= P(x(t_1)^T (X_3 \times ... \times X_{q+1})\, x(t_2)^T (X_3 \times ... \times X_{q+1}) < 0) - \frac{1}{2} \\
&\quad= P(\det(x(t_1), x(T_3), \dots, x(T_{q+1})) \cdot \det(x(t_2), x(T_3), \dots, x(T_{q+1})) < 0) - \frac{1}{2} \\
&\quad= P\Big(\prod_{j\geq 3}(T_j - t_1) \prod_{3\leq i<j\leq q+1}(T_j - T_i) \cdot \prod_{j\geq 3}(T_j - t_2) \prod_{3\leq i<j\leq q+1}(T_j - T_i) < 0\Big) - \frac{1}{2} \\
&\quad= P\Big(\prod_{j=3}^{q+1}(T_j - t_1)(T_j - t_2) < 0\Big) - \frac{1}{2} \\
&\quad= \frac{1}{2} - \frac{1}{2}(1 - 2P((T_1 - t_1)(T_1 - t_2) < 0))^{q-1} - \frac{1}{2} \\
&\quad= -\frac{1}{2}(1 - 2|F^{T_1}(t_1) - F^{T_1}(t_2)|)^{q-1}. \quad \square
\end{aligned}
$$

Proof of Proposition 3
At first we derive the Fourier series representation of f^r where f is given by

$$
f : [-1, 1] \ni z \longrightarrow f(z) = \frac{1}{2} - |z| \;\in \left[-\frac{1}{2}, \frac{1}{2}\right].
$$

Since

$$\left\{\frac{1}{\sqrt{2}},\ \cos(l\,\pi\cdot),\ \sin(l\,\pi\cdot);\ l\in I\!N\right\}$$

is an orthonormal basis of $I\!L_2[-1,1]$ and f^r is with f an even function, f^r can be represented only by $\frac{1}{\sqrt{2}}$ and the cosine functions, i.e.

$$f^r(z)=\alpha_0^{(r)}\cdot\frac{1}{\sqrt{2}}+\sum_{l=1}^{\infty}\alpha_l^{(r)}\cdot\cos(l\,\pi\,z).$$

Since f^r is continuous and piecewise differentiable, the series is uniformly convergent so that

$$\alpha_0^{(r)}=\int_{-1}^{1}f^r(z)\cdot\frac{1}{\sqrt{2}}\,dz\ =\ \sqrt{2}\int_0^1 f^r(z)\,dz$$

and for $l\geq 1$

$$\alpha_l^{(r)}=\int_{-1}^{1}f^r(z)\cdot\cos(l\,\pi\,z)\,dz=2\int_0^1 f^r(z)\cdot\cos(l\,\pi\,z)\,dz.$$

This implies for $r=1$

$$\alpha_0^{(1)}\ =\ 0,$$

and for $l\geq 1$

$$\alpha_l^{(1)}\ =\ 2\int_0^1\left(\frac{1}{2}-z\right)\cdot\cos(l\,\pi\,z)\,dz\ =\ \begin{cases}0, & \text{if } l \text{ is even,}\\ \frac{4}{l^2\,\pi^2}, & \text{if } l \text{ is odd.}\end{cases}\tag{10}$$

For $r=2$, we obtain

$$\alpha_0^{(2)}\ =\ 2\int_0^1\left(\frac{1}{2}-z\right)^2\cdot\frac{1}{\sqrt{2}}\,dz\ =\ \frac{\sqrt{2}}{12},$$

and for $l\geq 1$

$$\alpha_l^{(2)}\ =\ 2\int_0^1\left(\frac{1}{2}-z\right)^2\cdot\cos(l\,\pi\,z)\,dz\ =\ \begin{cases}\frac{4}{l^2\,\pi^2}, & \text{if } l \text{ is even,}\\ 0, & \text{if } l \text{ is odd.}\end{cases}\tag{11}$$

For $r>2$, we have

$$\alpha_0^{(r)}\ =\ \frac{2}{\sqrt{2}}\int_0^1\left(\frac{1}{2}-z\right)^r\,dz\ =\ \frac{2}{\sqrt{2}}\begin{cases}0, & \text{if } r \text{ is odd,}\\ \frac{1}{(r+1)\,2^r}, & \text{if } r \text{ is even,}\end{cases}$$

and for $l \geq 1$, partial integration provides the following recursion formula for $\alpha_l^{(r)}$

$$
\begin{aligned}
\alpha_l^{(r)} &= 2 \int_0^1 \left(\frac{1}{2} - z\right)^r \cdot \cos(l\,\pi\,z)\,dz \\
&= 2\,\frac{1}{l\,\pi}\,\sin(l\,\pi\,z)\,\left.\left(\frac{1}{2} - z\right)^r\right|_0^1 + \frac{2\,r}{l\,\pi}\int_0^1 \left(\frac{1}{2} - z\right)^{r-1} \cdot \sin(l\,\pi\,z)\,dz \\
&= \frac{2\,r}{l\,\pi}\int_0^1 \left(\frac{1}{2} - z\right)^{r-1} \cdot \sin(l\,\pi\,z)\,dz \\
&= -\frac{2\,r}{l\,\pi}\,\frac{1}{l\,\pi}\,\cos(l\,\pi\,z)\,\left.\left(\frac{1}{2} - z\right)^{r-1}\right|_0^1 - \frac{2\,r}{l\,\pi}\,\frac{(r-1)}{l\,\pi}\int_0^1 \left(\frac{1}{2} - z\right)^{r-2} \cdot \cos(l\,\pi\,z)\,dz \\
&= -\frac{2\,r}{l\,\pi}\,\frac{1}{l\,\pi}\left[(-1)^l\,\left(-\frac{1}{2}\right)^{r-1} - \left(\frac{1}{2}\right)^{r-1}\right] - \frac{2\,r}{l\,\pi}\,\frac{(r-1)}{l\,\pi}\,\frac{1}{2}\,\alpha_l^{(r-2)} \\
&= -\frac{r}{l^2\,\pi^2\,2^{r-2}}\,\left[(-1)^{l+r-1} - 1\right] - \frac{r\,(r-1)}{l^2\,\pi^2}\,\alpha_l^{(r-2)} \\
&= \begin{cases} -\frac{r\,(r-1)}{l^2\,\pi^2}\,\alpha_l^{(r-2)}, & \text{if } l + r \text{ is odd,} \\ \frac{r}{l^2\,\pi^2}\,\left[\frac{1}{2^{r-3}} - (r-1)\,\alpha_l^{(r-2)}\right], & \text{if } l + r \text{ is even.} \end{cases}
\end{aligned}
$$

Since $\alpha_l^{(1)} = 0$ if l is even and $\alpha_l^{(2)} = 0$ if l is odd, we obtain $\alpha_l^{(r)} = 0$ if $r + l$ is odd. If $r + l$ is even and $l \geq 1$, then we have

$$
\alpha_l^{(r)} = - \sum_{\substack{k \in \{1,\ldots,r\} \\ k \text{ odd}}} \frac{r!}{2^{r-k-2}(r-k)!}(-l^2\,\pi^2)^{-\frac{k+1}{2}}.
$$

This can be seen by induction over r: for $r = 1$ and l odd, it holds according to (10)

$$
- \sum_{\substack{k \in \{1,\ldots,r\} \\ k \text{ odd}}} \frac{r!}{2^{r-k-2}(r-k)!}(-l^2\,\pi^2)^{-\frac{k+1}{2}} = -\frac{1!}{2^{-2}\,0!}(-l^2\,\pi^2)^{-1} = \frac{4}{l^2\,\pi^2} = \alpha_l^{(1)},
$$

and for $r = 2$ and l even, it holds according to (11)

$$
- \sum_{\substack{k \in \{1,\ldots,r\} \\ k \text{ odd}}} \frac{r!}{2^{r-k-2}(r-k)!}(-l^2\,\pi^2)^{-\frac{k+1}{2}} = -\frac{2!}{2^{-1}\,1!}(-l^2\,\pi^2)^{-1} = \frac{4}{l^2\,\pi^2} = \alpha_l^{(2)}.
$$

The induction step is done from r to $r+2$, that is:

$$\begin{aligned}
\alpha_l^{(r+2)} &= \frac{r+2}{l^2\,\pi^2}\left[\frac{1}{2^{r-1}} - (r+1)\,\alpha_l^{(r)}\right] \\
&= \frac{r+2}{l^2\,\pi^2}\left[\frac{1}{2^{r-1}} + (r+1)\sum_{\substack{k\in\{1,\dots,r\}\\ k \text{ odd}}} \frac{r!}{2^{r-k-2}(r-k)!}(-l^2\,\pi^2)^{-\frac{k+1}{2}}\right] \\
&= \frac{(r+2)!}{2^{r+2-3}\,(r+2-1)!}(l^2\,\pi^2)^{-1} \\
&\qquad - \sum_{\substack{k\in\{1,\dots,r\}\\ k \text{ odd}}} \frac{(r+2)!}{2^{r+2-(k+2)-2}(r+2-(k+2))!}(-l^2\,\pi^2)^{-\frac{k+2+1}{2}} \\
&= -\sum_{\substack{k\in\{1,\dots,r+2\}\\ k \text{ odd}}} \frac{(r+2)!}{2^{r+2-k-2}(r+2-k)!}(-l^2\,\pi^2)^{-\frac{k+1}{2}}.
\end{aligned}$$

Hence, we always have $\alpha_0^{(r)} = \sqrt{2}\,\gamma_0^{(r)}$ and $\alpha_l^{(r)} = 2\,\gamma_l^{(r)}$ for $l \geq 1$, where $\gamma_l^{(r)}$ are the quantities of Theorem 3.

To finish the proof, we transfer the Fourier series representation of $f^r(z)$ on $[-1,1]$ to that of $g^r(s,t) = f^r(s-t)$ on $[0,1]^2$. This provides

$$\begin{aligned}
\left(\frac{1}{2} - |s-t|\right)^r &= f^r(s-t) = \alpha_0^{(r)}\cdot\frac{1}{\sqrt{2}} + \sum_{l=1}^{\infty}\alpha_l^{(r)}\cdot\cos(l\,\pi\,(s-t)) \\
&= \alpha_0^{(r)}\cdot\frac{1}{\sqrt{2}} + \sum_{l=1}^{\infty}\alpha_l^{(r)}\cdot[\cos(l\,\pi\,s)\cdot\cos(l\,\pi\,t) + \sin(l\,\pi\,s)\cdot\sin(l\,\pi\,t)]
\end{aligned}$$

which is the representation given by Theorem 3 using the relation between $\alpha_l^{(r)}$ and $\gamma_l^{(r)}$. The quantities $\gamma_l^{(r)}$ are used in Theorem 3 since only

$$\mathcal{S} = \left\{1,\ \sqrt{2}\cos(l\,\pi\cdot),\ \sqrt{2}\sin(l\,\pi\cdot);\ l\in I\!N\right\}$$

are normalized functions of $I\!L_2[0,1]$. However, $\mathcal{S}$ is not an orthonormal system of $I\!L_2[0,1]$. But, since the quantities $\gamma_l^{(r)}$ are zero as soon as $r+l$ is odd, only the systems

$$\begin{aligned}
&\left\{\sqrt{2}\cos(l\,\pi\cdot),\ \sqrt{2}\sin(l\,\pi\cdot);\ l\in I\!N \text{ and } l \text{ is odd}\right\} \text{ for } r \text{ odd}, \\
&\left\{1,\ \sqrt{2}\cos(l\,\pi\cdot),\ \sqrt{2}\sin(l\,\pi\cdot);\ l\in I\!N \text{ and } l \text{ is even}\right\} \text{ for } r \text{ even},
\end{aligned}$$

are relevant and these are orthonormal systems of $I\!L_2[0,1]$. □

Acknowledgement

The data base for the example is created by Stanislav Katina. We thank him for permission to include them.

References

[1] Arcones, M.A., Chen, Z. and Giné, E. (1994). Estimators related to U-processes with applications to multivariate medians: asymptotic normality. *Ann. Statist.* **22**, 1460-1477.

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

[3] Christmann, A. and Rousseeuw, P.J. (2001). Measuring overlap in logistic regression. *Computational Statistics and Data Analysis* **28**, 65-75.

[4] Dümbgen, L. (1992). Limit theorems for the simplicial depth. *Statist. Probab. Lett.* **14**, 119-128.

[5] Lee, A.J. (1990). *U-Statistics. Theory and Practice.* Marcel Dekker, New York.

[6] Liu, R.Y. (1988). On a notion of simplicial depth. *Proc. Nat. Acad. Sci. USA* **85**, 1732-1734.

[7] Liu, R.Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18**, 405-414.

[8] Liu, R.Y. (1992). Data depth and multivariate rank tests. In *L_1-Statistical Analysis and Related Methods*, ed. Y. Dodge, North-Holland, Amsterdam, 279-294.

[9] 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.

[10] Mizera, I. (2002). On depth and deep points: A calculus. *Ann. Statist.* **30**, 1681-1736.

[11] Mizera, I. and Müller, Ch.H. (2004). Location-scale depth. *Journal of the American Statistical Association* **99**, 949-966. With discussion.

[12] Mosler, K. (2002). *Multivariate Dispersion, Central Regions and Depth. The Lift Zonoid Approach.* Lecture Notes in Statistics **165**, Springer, New York.

[13] Müller, Ch.H. (2005). Depth estimators and tests based on the likelihood principle with application to regression. *Journal of Multivariate Analysis* **95**, 153-181.

[14] Rousseeuw, P.J. and Hubert, M. (1999). Regression depth (with discussion). *J. Amer. Statist. Assoc.* **94**, 388-433.

[15] Tricomi, F.G. (1955). *Vorlesungen über Orthogonalreihen.* Springer, Berlin.

[16] Tukey, J.W. (1975). Mathematics and the picturing of data. In *Proc. International Congress of Mathematicians*, Vancouver 1974, **2**, 523- 531.

[17] Van Aelst, S., Rousseeuw, P.J., Hubert, M. and Struyf, A. (2002). The deepest regression method. *J. Multivariate Anal.* **81**, 138-166.

[18] Wellmann, R., Katina, S. and Müller, Ch.H. (2007a). Calculation of simplicial depth estimators for polynomial regression with applications. *Computational Statistics and Data Analysis* **51**, 5025-5040.

[19] Wellmann, R. and Müller, Ch.H. (2007b). Tests for multiple regression based on simplicial depth. *Submitted.*

[20] Wellmann, R. (2007c). On data depth with application to regression and tests. *Ph.D. thesis.* University of Kassel, Germany.

[21] Witting, H. and Müller-Funk, U. (1995). *Mathematische Statistik II.* Teubner, Stuttgart.

[22] Zuo, Y. and Serfling, R. (2000a). General notions of statistical depth function. *Ann. Statist.* **28**, 461-482.

[23] Zuo, Y. and Serfling, R. (2000b). Structural properties and convergence results for contours of sample statistical depth functions. *Ann. Statist.* **28**, 483-499.

wellmann@mathematik.uni-kassel.de
Department of Mathematics
University of Kassel
D-34109 Kassel
Germany