KYBERNETIKA — VOLUME *49* (2013), NUMBER 4, PAGES 644–662

REFERENCE POINTS BASED TRANSFORMATION AND APPROXIMATION

Csaba Török

Interpolating and approximating polynomials have been living separately more than two centuries. Our aim is to propose a general parametric regression model that incorporates both interpolation and approximation.

The paper introduces first a new r-point transformation that yields a function with a simpler geometrical structure than the original function. It uses $r \geq 2$ reference points and decreases the polynomial degree by $r-1$. Then a general representation of polynomials is proposed based on $r \geq 1$ reference points.

The two-part model, which is suited to piecewise approximation, consist of an ordinary least squares polynomial regression and a reparameterized one. The later is the central component where the key role is played by the reference points. It is constructed based on the proposed representation of polynomials that is derived using the r-point transformation $T_r(x)$. The resulting polynomial passes through r reference points and the other points approximates. Appropriately chosen reference points ensure quasi smooth transition between the two components and decrease the dimension of the LS normal matrix. We show that the model provides estimates with such statistical properties as consistency and asymptotic normality.

Keywords: polynomial representation, approximation model, smooth connection, consistency, asymptotic normality

Classification: 41A10, 65D05, 65D07, 65D10, 62J05, 62F12

1. INTRODUCTION

Consider two polynomials $P_{\mathbf{a}}(x)$ and $P_{\mathbf{b}}(x)$ over the intervals $[-1, 0]$ and $[0, 1]$, respectively, where $\mathbf{a} = (a_0, a_1, \ldots, a_p)^T$, $\mathbf{b} = (b_0, b_1, \ldots, b_q)^T$, the degrees p, q are finite and $p,\ q \geq r-1 \geq 0$. To guarantee a *smooth* transition between the polynomials, we will demand that the following *quasi spline* conditions to order $r-1$ hold in the shared abscissa zero

$$\begin{aligned} P_{\mathbf{a}}(0) &= P_{\mathbf{b}}(0), \\ P_{\mathbf{a}}^{(j)}(0) &= P_{\mathbf{b}}^{(j)}(0) + o(\tau), \quad j = \overline{1,\, r-1},\ \ r \geq 2, \end{aligned} \tag{1}$$

i. e. instead of the spline conditions $P_{\mathbf{a}}^{(j)}(0) = P_{\mathbf{b}}^{(j)}(0)$ we require only that the derivatives in the shared point should be sufficiently close to each other $|P_{\mathbf{a}}^{(j)}(0) - P_{\mathbf{b}}^{(j)}(0)| < c_j\tau$, where τ is a small positive real number and $P^{(j)}$ denotes the jth derivative of P. We will not use in our models derivatives however it is well known that these conditions can be achieved by finite differences using nearby function values. Naturally, if $r = 1$, then we can guarantee only continuity in zero. For each M and N we have a set of observations from $[-1,\, 0]$ and $[0,\, 1]$

$$\{[x_{i,M}, \tilde{y}_{i,M}],\ x_{i,M} = -\frac{i}{M},\ \tilde{y}_{i,M} = P_{\mathbf{a}}(x_{i,M}) + \varepsilon^*_{i,M},\ i = \overline{1, M}\},$$
$$\{[x_{j,N}, \tilde{y}_{j,N}],\ x_{j,N} = \frac{j}{N},\ \tilde{y}_{j,N} = P_{\mathbf{b}}(x_{j,N}) + \varepsilon_{j,N},\ j = \overline{1, N}\}.$$

We will assume that $\{\varepsilon^*_{i,M}\}_{i=\overline{1,M}}$ and $\{\varepsilon_{j,N}\}_{j=\overline{1,N}}$ are uncorrelated error sequences, $E\varepsilon^*_{i,M} = 0,\ E\varepsilon_{j,N} = 0,\ E\varepsilon^*_{i,M}\varepsilon^*_{I,M} = \sigma^2_*\delta_{i,I} < \infty,\ E\varepsilon_{j,N}\varepsilon_{J,N} = \sigma^2\delta_{j,J} < \infty$ and $E\varepsilon^*_{i,M}\varepsilon_{j,N} = 0,\ \ i, I = \overline{1, M},\ j, J = \overline{1, N}$, where $\delta_{\cdot,\cdot}$ is the Kronecker's delta, and the integers $M, N \gg 1,\ M = \kappa N$, where κ is any positive rational number such that κN is integer.

Our aim is to find approximants for the polynomials with the foregoing smooth conditions. It must be underlined that we will solve the question of smoothness of the two polynomials in their common point differently from splines thanks to reference points.

Let us introduce a set of reference points with unequal abscissas

$$\mathcal{R} \equiv \mathcal{R}_{\mathbf{b}} = \{\, [v_j,\, y_j],\ \ y_j = P_{\mathbf{b}}(v_j),\ \ v_i \neq v_j \text{ for } i \neq j,\ \ i, j = \overline{0, r-1}\ \}.$$

We will investigate a two-part model

$$\begin{array}{rcllcl} \tilde{y}^*_{i,M} &=& P_{\mathbf{a}}(-i/M) + \varepsilon^*_{i,M}, & i &=& \overline{1,\, M}, \\ \tilde{y}_{j,N} &=& I_{\mathcal{R}}(j/N) + \mathbf{w}_r^T(j/N)\cdot\boldsymbol{\beta} + \varepsilon_{j,N}, & j &=& \overline{1,\, N}, \end{array} \tag{2}$$

where $I_{\mathcal{R}}(x)$ is an interpolation polynomial of order $r-1$ defined by the r reference points from $\mathcal{R}$, $\boldsymbol{\beta} = (b_r, b_{r+1}, \ldots, b_q)^T$ and the vector of the basis functions $\mathbf{w}_r$ is constructed using the abscissas $v_0,\ v_1 \ldots, v_{r-1}$ of the reference points. We emphasize that due to reparameterization the second part of the model does not contain the coefficients $b_0,\ b_1 \ldots, b_{r-1}$, and $I_{\mathcal{R}}(x)$ and $\mathbf{w}_r(x)$ do not depend on $b_r,\ b_{r+1} \ldots, b_q$.

The model of the first part can be a classical polynomial regression. The central component of the two-part model is the second model. Here the key role is played by the reference points. The paper shows that their appropriate localization assures smooth transition between the polynomials and in the successive mode proposed estimate $\widehat{\boldsymbol{\beta}}_N$ is consistent and $\sqrt{N}(\widehat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$ is asymptotically normal.

Interpolating and approximating regression polynomials have been living alongside more than two centuries [8]. Both of them have diverse philosophy and their own typical advantages and drawbacks [8, 15], so it is a real challenge to search a way to combine them to leverage their benefits. The paper proposes a new general way to achieve this goal.

The nature of dependence between variables and the level of the accuracy of data points define the tools and techniques with a given criteria of the goodness-of-fit the researchers can use. While interpolation is suited to approximate errorless data, regression analysis is an optimal tool for approximation, smoothing noisy data [15]. Noisy data with complex structure can be approximated using such nonparametric methods as smoothing splines [7], kernel smoothers [11] with loose conditions [1] or neural networks [14]. In the present paper developed approach to smoothing noisy data with complex dependence is based on the two-part parametric model (2) with a decreased number of estimated parameters. Our Clenshaw-Hayes type [15] model that is based on the IZA (Interpolation, Zero and Approximation) representation incorporates both interpolation and approximation. It was derived using the proposed r-point transformation $T_r(x)$ that was motivated by the discrete projective transformation (DPT) introduced by N.D.Dikoussar [2]. DPT can be obtained from $T_r(x)$ with $r = 3$.

DPT was applied successfully to various approximation problems. The papers [10, 16] showed how it can be used to assess the unknown degree in regression polynomials. [17] was devoted to local function approximation based on the inverse discrete projective transformation. The modification of the technique used there led to a three-point cubic model that happened to be very useful in local approximation and allowed us to develop a global autotracking piecewise cubic approximation for data with moderate error [3, 5, 12]. Our first results in approximation of 3D data and signal compression based on the three-point approach are contained in [6] and [9], respectively. However we did not give a satisfactory answer in these works to the two base questions of the piecewise three-point approach: how to assess the ordinates of the reference points and how to guarantee a smooth transition between the neighboring polynomials and their approximants. These questions for recursive approximation with cubic polynomials was answered in [13], where the asymptotic normality of the estimate was also proved. For solving the above issues of non recursive polynomial approximation the present paper generalizes the two-part approach of [13] based on the IZA representation of polynomials. Dikoussar enhanced his tree point representation in another direction resulted in the basic element method [4].

The structure of the paper is as follows. The road to the fusion of interpolating and approximating polynomials can lead through the r-point transformation defined in Section 2. The section shows that the transformation decreases the degree of polynomials by $r-1$. The general formula for IZA polynomials is derived in Section 3. The consecutive section includes the main results about the model and its properties.

2. THE R-POINT TRANSFORMATION

In this section first we introduce a new r-point transformation of functions and its inverse, and then derive a formula for the transformation of polynomials that will be used in the next section.

Definition 2.1. The forward r-point transformation, $r-1 \in \mathbb{Z}^+$, of any sufficiently smooth function $f(x)$ based on a set of r reference points $\mathcal{R} = \{[v_i, y_i],\ y_i = f(v_i),\ i = \overline{0, r-1}\}$ is given by

$$f^T(x) \equiv T_{\mathcal{R}} f(x) = H_0(x) f(x) + \sum_{i=1}^{r-1} H_i(x) y_i, \tag{3}$$

where

$$H_0(x) = \frac{1}{\Pi_0(x)}, \quad H_i(x) = -\frac{\Pi_i(x)}{\Pi_0(x)}, \;\; i = \overline{1, r-1},$$

$$\Pi_i(x) = \prod_{\substack{0 \le k \le r-1 \\ k \ne i}} \frac{x - v_k}{v_i - v_k}, \;\; i = \overline{0, r-1}, \tag{4}$$

$v_0 \neq v_1 \neq \ldots \neq v_{r-1}$ and $x \neq v_i, \; i = \overline{1, r-1}$.

Notice that the transformation does not use y_0. For the sake of brevity, instead of $T_{\mathcal{R}} f(x)$ we will usually write $T_r f(x), \; r \geq 2$. We mention that Π_i are the well known parabolas from the the Waring–Euler–Lagrange interpolating polynomial $I_{r-1}(x) = \sum_{i=0}^{r-1} \Pi_i(x) y_i$, and the transformation written explicitly

$$\begin{aligned} T_r f(x) \;\; &= \;\; \overbrace{\frac{(v_0 - v_1)(v_0 - v_2)\ldots(v_0 - v_{r-1})}{(x - v_1)(x - v_2)\ldots(x - v_{r-1})}}^{H_0(x)} f(x) + \\ &+ \underbrace{\frac{(v_0 - x)(v_0 - v_2)\ldots(v_0 - v_{r-1})}{(v_1 - x)(v_1 - v_2)\ldots(v_1 - v_{r-1})}}_{H_1(x)} y_1 + \ldots + \underbrace{\frac{(v_0 - x)(v_0 - v_1)\ldots(v_0 - v_{r-2})}{(v_{r-1} - x)(v_{r-1} - v_1)\ldots(v_{r-1} - v_{r-2})}}_{H_{r-1}(x)} y_{r-1}. \end{aligned}$$

So to transform a function $y = f(x)$ or a point $[x, f(x)]$ by T_r, one needs $r-1$ points $[v_1, y_1], [v_2, y_2], \ldots, [v_{r-1}, y_{r-1}]$ from the function $f(x)$ and another value v_0. The only assumption about $x, \; v_0, \ldots, v_{r-1}$ is that they should be mutually different except for x and v_0 (see definition 2.1). The reference point $[v_0, y_0]$ is a special one: its second coordinate appears not in the definition of the transformation but in its result, see lemma 2.4 and theorem 2.8. It is a fix point as the transformation does not change it, see lemma 2.3.

We obtain the backward r-point transformation simply expressing $f(x)$ from (3).

Lemma 2.2. The backward r-point transformation is given by

$$T_{\mathcal{R}}^{-1} f^T(x) \equiv T_{\mathcal{R}}^{-1} T_{\mathcal{R}} f(x) = \Pi_0(x) T_{\mathcal{R}} f(x) + \sum_{i=1}^{r-1} \Pi_i(x) y_i,$$

where $\Pi_i(x), \; i = \overline{0, r-1}$ are defined by (4).

The index $r-1$ in $I_{r-1}(x)$ indicates the polynomial's degree while r the count of the reference points.

The inverse transformation T_r^{-1} written explicitly

$$\begin{aligned} T_r^{-1} T_r f(x) \;\; &= \;\; \overbrace{\frac{(x - v_1)(x - v_2)\ldots(x - v_{r-1})}{(v_0 - v_1)(v_0 - v_2)\ldots(v_0 - v_{r-1})}}^{\Pi_0(x)} T_r f(x) + \\ &+ \underbrace{\frac{(x - v_0)(x - v_2)\ldots(x - v_{r-1})}{(v_1 - v_0)(v_1 - v_2)\ldots(v_1 - v_{r-1})}}_{\Pi_1(x)} y_1 + \ldots + \underbrace{\frac{(x - v_0)(x - v_1)\ldots(x - v_{r-2})}{(v_{r-1} - v_0)(v_{r-1} - v_1)\ldots(v_{r-1} - v_{r-2})}}_{\Pi_{r-1}(x)} y_{r-1}. \end{aligned}$$

We recommend to compare for the simplest case $r = 2$ the transformation $T_2 f(x)$, its inverse $T_2^{-1} T_2 f(x)$ and the linear interpolation $I_1(x)$, that use $[v_0, y_0]$, $[v_1, y_1]$, $y_i = f(v_i)$, $i = 0, 1$:

$$\begin{aligned} T_2 f(x) &= \frac{v_0 - v_1}{x - v_1} f(x) + \frac{v_0 - x}{v_1 - x} y_1, \\ T_2^{-1} T_2 f(x) &= \frac{x - v_1}{v_0 - v_1} T_2 f(x) + \frac{x - v_0}{v_1 - v_0} y_1, \\ I_1(x) &= \frac{x - v_1}{v_0 - v_1} y_0 + \frac{x - v_0}{v_1 - v_0} y_1. \end{aligned}$$

We mention that from the well known relation $\sum_{i=0}^{r-1} \Pi_i(x) = 1$ it follows

$$\sum_{i=0}^{r-1} H_i(x) = 1. \tag{5}$$

Lemma 2.3.
a) The r-point transformation has the property of linearity, i. e. for any sufficiently smooth function $f(x)$ and real c_1, c_2

$$T_r(c_1 f(x) + c_2) = c_1 T_r f(x) + c_2,$$

b) for the T_r transformation $[v_0, f(v_0)]$ is a fix point

$$T_{\mathcal{R}_r} f(v_0) = f(v_0) = T_{\mathcal{R}_s} f(v_0), \;\; r \neq s.$$

P r o o f. a) From the definition 2.1 we get

$$\begin{aligned} T_r(c_1 f(x) + c_2) &= H_0(x)(c_1 f(x) + c_2) + \sum_{i=1}^{r-1} H_i(x)(c_1 f(v_i) + c_2) \\ &= c_1 \Big(H_0(x) f(x) + \sum_{i=1}^{r-1} H_i(x) f(v_i) \Big) + c_2 \sum_{i=0}^{r-1} H_i(x), \end{aligned}$$

and it remained to apply the relation (5).

The proof of the part b) follows directly from the definition of T_r. □

The properties of DPT are studied in most detail for power functions and polynomials [2, 16]. It is known that the DPT decreases the polynomial's degree by two. We generalize this result. First we deal with power functions. The next lemma given without proof showes that the r-point transformation T_r maps a power function of degree p to a polynomial function of degree $p - r + 1$.

Lemma 2.4. Let $r - 1 \in \mathbb{Z}^+$ and $p \in \mathbb{Z}^{0+}$. The r-point transformation of x^p is
a) a constant for $p < r$

$$T_r x^p = v_0^p, \tag{6}$$

b) a polynomial of degree $p-(r-1)$ for $p \geq r$

$$T_r x^p = v_0^p + z_1\ S_r^{p-r}, \tag{7}$$

where

$$z_1 \equiv z_1(x) = (x - v_0) \prod_{i=1}^{r-1} (v_0 - v_i), \tag{8}$$

$$S_r^{p-r}(x) \equiv S_r^{p-r} = \sum_{k_1=0}^{k_0} \sum_{k_2=0}^{k_1} \dots \sum_{k_{p-r}=0}^{k_{p-r-1}} v_{k_1} v_{k_2} \dots v_{k_{p-r}}, \quad k_0 \equiv r,\ r < p \tag{9}$$

$v_r \equiv x$ and $S_r^0 = 1$.

Before giving a recursive formula for evaluating S_r^j, and examples of both S_r^j and $T_r x^p$, we make some remarks.

z_1 and S_*^{p-r} are polynomials of degree 1 and $p-r$, respectively. Notice different use of $y_0 \equiv v_0^p$ in (6) and (7). In (7) the use of v_0^p enables the factorization, however it is canceled after performing the given operations.

The symmetrical polynomials $S_r^j,\ j \geq 0,\ r \geq 2$, play a fundamental role in expressing the result of the transformation of power functions. The upper index j indicates the maximal power of v_i in S_r^j and the lower one r reminds that it contains in addition to the abscissas $v_0, \dots v_{r-1}$ of the reference points as well as $x \equiv v_r$. More precisely

$$S_r^j = \sum_{\substack{m_0,\ m_1,\ \dots,\ m_r\ =\ 0 \\ m_0\ +\ m_1\ +\ \dots\ +\ m_r\ =\ j}} v_0^{m_0} v_1^{m_1} \dots v_r^{m_r}.$$

Now let us see a simple relation based on which S_r^j can be evaluated recursively (even for $r = 1$).

Lemma 2.5. Let $S_r^0 = 1,\ r \geq 0$, and $S_0^j = v_0^j,\ j \geq 0$. Then for any $j \geq 1,\ r \geq 1$

$$S_r^j = \sum_{i=0}^{r} v_i S_i^{j-1}, \tag{10}$$

$$= S_{r-1}^j + v_r S_r^{j-1}. \tag{11}$$

Proof. From (9) we get $S_r^{p-r} = \sum_{i=0}^r v_i\ S_i^{p-r-1}$, and hence

$$S_r^j = \sum_{i=0}^{r} v_i S_i^{j-1} = S_{r-1}^j + v_r S_r^{j-1}.$$

□

Example 2.6. Now let us see some examples of S_r^j and $T_r x^p$. First of all we mention that for any $r \geq 1$

$$S_r^1 = \sum_{i=0}^{r} v_i.$$

We apply (11) for computing S_R^J successively until we get 1 for J and 0 for R, and at the end we leverage the relations $S_0^j = v_0^j$, $S_r^1 = \sum_{i=0}^r v_i$, and make the substitution $v_r \equiv x$, $r \geq 2$.

Since $S_1^2 = S_0^2 + v_1 S_1^1 = v_0^2 + v_1(v_0 + v_1)$, we get

$$\begin{aligned} S_2^3 &= S_1^3 + v_2 S_2^2 = (S_0^3 + v_1 S_1^2) + v_2(S_1^2 + v_2 S_2^1) \\ &= v_0^3 + v_1(v_0^2 + v_1(v_0 + v_1)) + x(v_0^2 + v_1(v_0 + v_1) + x(v_0 + v_1 + x)). \end{aligned}$$

Based on (10) the computation is as follows

$$\begin{aligned} S_2^2 &= v_0 S_0^1 + v_1 S_1^1 + v_2 S_2^1 \\ &= v_0^2 + v_1(v_0 + v_1) + x(v_0 + v_1 + x) \end{aligned}$$

and

$$\begin{aligned} S_2^3 &= v_0 S_0^2 + v_1 S_1^2 + v_2 S_2^2 \\ &= v_0^3 + v_1(v_0^2 + v_1(v_0 + v_1)) + x(v_0^2 + v_1(v_0 + v_1) + x(v_0 + v_1 + x)). \end{aligned}$$

From (7), (11) we obtain

$$\begin{aligned} T_2 x^4 &= {v_0}^4 + (v_0 - v_1)(x - v_0)\left(x^2 + (v_0 + v_1)\,x + {v_0}^2 + v_1\,(v_0 + v_1)\right), \\ T_3 x^3 &= {v_0}^3 + (v_0 - v_1)(v_0 - v_2)(x - v_0), \\ T_4 x^5 &= {v_0}^5 + (v_0 - v_1)(v_0 - v_2)(v_0 - v_3)(x - v_0)(x + v_0 + v_1 + v_2 + v_3). \end{aligned}$$

The next useful relation, obtained by multiple application of (11), contains all exponents of $x \equiv v_r$ explicitly.

Corollary 2.7. For any $J \geq 1, \; r \geq 1$

$$S_r^J = \sum_{j=0}^{J} x^{J-j} S_{r-1}^j.$$

Lemma 2.4 and the linearity of T_r imply that T_r decreases the degree of a polynomial function

$$P_p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_p x^p, \quad p \geq 0,$$

by $r-1$. The exact formulae are given by

Theorem 2.8. Let $r - 1 \in \mathbb{Z}^+$ and $p \in \mathbb{Z}^{0+}$. The r-point transformation of $P_p(x)$ is
a) a constant for $p < r$

$$T_r P_p(x) = P_p(v_0),$$

b) a polynomial of degree $p - (r-1)$ for $p \geq r$

$$T_r P_p(x) = P_p(v_0) + z_1\, A_{p-r,r}(x),$$

where z_1 is defined by (8),

$$A_{p-r,r}(x) = \mathbf{S}^T \cdot \boldsymbol{\alpha}, \tag{12}$$

$\mathbf{S} = (S_r^0,\ S_r^1\ \ldots,\ S_r^{p-r})^T$ and $\boldsymbol{\alpha} = (a_r,\ a_{r+1}\ \ldots,\ a_p)^T$.

Before proving this result we mention that since the r-point transformation T_r decreases the polynomial degree by $r-1$, the coefficients $a_i,\ i=\overline{0,r-1}$, are absent from the result of $T_r P_p(x)$ and consequently $A_{p-r,r}$ contains only the last $p-r+1$ coefficients $a_i,\ i=\overline{r,p}$.

P r o o f. b) Based on the definition and linearity of the T_r transformation, and lemma 2.4 we get

$$\begin{aligned} T_r P_p(x) &= T_r \sum_{i=0}^{r-1} a_i x^i + T_r \sum_{i=r}^{p} a_i x^i = \sum_{i=0}^{r-1} a_i T_r x^i + \sum_{i=r}^{p} a_i T_r x^i \\ &= \sum_{i=0}^{r-1} a_i v_0^i + \sum_{i=r}^{p} a_i v_0^i + \sum_{i=r}^{p} a_i z_1 S_r^{i-r} \\ &= P_p(v_0) + z_1\, \mathbf{S}^T \cdot \boldsymbol{\alpha}. \end{aligned}$$

□

Example 2.9. Let us see some instances of $T_r P_p(x)$ and $A_{t,r},\ t=p-r$.

$$\begin{aligned} T_2 P_4(x) &= P_4(v_0) + (v_0 - v_1)\,(x - v_0)\,A_{2,2}, \\ T_3 P_5(x) &= P_5(v_0) + (v_0 - v_1)\,(v_0 - v_2)\,(x - v_0)\,A_{2,3}, \end{aligned}$$

where

$$\begin{aligned} A_{2,2} &= a_2 + a_3\,(x + v_0 + v_1) + a_4\Big(x^2 + (v_0 + v_1)\,x + {v_0}^2 + v_1\,(v_0 + v_1)\Big), \\ A_{2,3} &= a_3 + a_4\,(x + v_0 + v_1 + v_2) \\ &\quad + a_5\Big(x^2 + (v_0 + v_1 + v_2)\,x + {v_0}^2 + v_1\,(v_0 + v_1) + v_2\,(v_0 + v_1 + v_2)\Big). \end{aligned}$$

Since the relation between the polynomial degree p and the reference points' count r, as well as their minimal values are changing in the paper's sections, table 1 gives a short preview of them for the main concepts.

$T_r P_p(x)$ transformation	$p+1 \geq r \geq 2$
IZA(p,r) representation	$p+1 \geq r \geq 1$
IZA model	$p+1 \geq r \geq 1$
Estimator normality	$p+1 > r \geq 1$

Tab. 1. Relations between p and r, and their minimal values.

3. IZA POLYNOMIALS

This section is devoted to a general representation of polynomials and the next one to the approximation model.

Theorem 3.1. Assume $p \geq r-1 \geq 0$. Then any polynomial $P_p(x)$ can be expressed based on its any different r points $\{[v_i,\, y_i],\ y_i = P_p(v_i),\ i=\overline{0,r-1}\}$ as

$$P_p(x) = I_{r-1}(x) + Z_r(x) A_{p-r,r}(x), \tag{13}$$

where

$$I_{r-1}(x) = \sum_{i=0}^{r-1} \Pi_i(x) y_i$$

is an incomplete, if $p > r-1$, interpolating polynomial, $\Pi_i(x)$ and $A_{p-r,r}(x)$ are defined by (4) and (12), $A_{-1,r}(x) \equiv 0$, and

$$Z_r(x) = \prod_{i=0}^{r-1} (x - v_i). \tag{14}$$

One can verify that after substituting $\sum_{j=0}^{p} a_j v_i^j$ into y_i, $i = \overline{0, r-1}$, in $I_{r-1}(x) = \sum_{i=0}^{r-1} \Pi_i(x) y_i$, and algebraic simplification of (13) we get $a_0 + a_1 x + a_2 x^2 + \cdots + a_p x^p$, nevertheless we give a simple derivation of the representation formula (13) based on the r-point forward and backward transformation of polynomials. However before the proof some remarks.

We call the relation $P = I + ZA$ from (13) the **IZA**(p, r) **representation of a polynomial** or merely IZA polynomial. It simply gives a formula for a polynomial of degree p that passes through $r \geq 1$ points. The polynomials $I_{r-1}(x)$, $Z_r(x)$ and $A_{p-r,r}(x)$ are of degree $r-1$, r and $p-r$, respectively, and at the r arguments v_i, $i = \overline{0, r-1}$,

$$I_{r-1}(v_i) = P_p(v_i) \qquad \text{and} \qquad Z_r(v_i) = 0,$$

i.e. whereas at every v_i the polynomial I_{r-1} passes through the polynomial P_p, Z_r is zero. The IZA reprezentation includes two particular cases: IZA$(r-1, r)$, when $A_{p-r,r}(x) = 0$ and IZA$(p, 1)$, when $I_0(x) = \Pi_0(x) \equiv P_p(v_0)$. The former gives the Waring-Euler-Lagrange interpolation formula and the latter the Bézuet's theorem on the polynomial remainder with an explicit expression for the quotient $A(p-1, 1)$. For example in the case of IZA(4,1) we have

$$P_4(x) = P_4(v_0) + (x - v_0)\left(a_1 + (x + v_0)\, a_2 + \left(x^2 + x v_0 + {v_0}^2\right) a_3 + \left(x^3 + v_0 x^2 + x {v_0}^2 + {v_0}^3\right) a_4\right).$$

P r o o f. It is sufficient to consider the case, when $p > r-1 > 0$, since $p = r-1$ and $r = 1$ give the above mentioned two particular cases. Let G denote the second part of the backward transformation, $G(x) = \sum_{i=1}^{r-1} \Pi_i(x) P_p(v_i)$. Based on the definition of T_r^{-1} and theorem 2.8 we get (notice that in the proof $[v_0, P_p(v_0)]$ becomes the fix point of T_r):

$$\begin{aligned} P_p(x) &= T_r^{-1}(T_r P_p(x)) = \Pi_0(x) T_r P_p(x) + G(x) \\ &= \Pi_0(x) P_p(v_0) + \Pi_0(x) z_1 A_{p-r,r}(x) + G(x) \\ &= I_{r-1}(x) + \prod_{k=0}^{r-1} (x - v_k) A_{p-r,r}(x), \end{aligned}$$

since from (4), (8) we obtain

$$\Pi_0(x) z_1 = \prod_{\substack{0 \leq k \leq r-1 \\ k \neq 0}} \frac{x - v_k}{v_0 - v_k} \prod_{i=0}^{r-1} (v_0 - v_i) = \prod_{k=0}^{r-1} (x - v_k).$$

□

We emphasize that from the $p+1$ coefficients $a_0, a_1, \ldots, a_p$ *of* $P_p(x)$ the IZA representation explicitly contains due to the reparameterization only the last $p-r+1$ ones within $A_{p-r,r}(x)$.

Example 3.2. We illustrate the IZA(p,r) representation of polynomials of degree p with r reference points

$$P_p(x) = I_{r-1}(x) + Z_r(x) A_{p-r,r}(x)$$

by examples for $p \geq r$.

`IZA(4,2):`
$$P_4(x) = \frac{(x-v_1)\,P_4(v_0)}{v_0-v_1} + \frac{(x-v_0)\,P_4(v_1)}{v_1-v_0} + (x-v_0)\,(x-v_1)$$
$$\Big(a_2 + a_3 \underbrace{(x+v_0+v_1)}_{S_2^1} + a_4 \underbrace{(x^2+(v_0+v_1)\,x+{v_0}^2+v_1\,(v_0+v_1))}_{S_2^2}\Big).$$

`IZA(3,3):`
$$P_3(x) = \frac{(x-v_1)\,(x-v_2)\,P_3(v_0)}{(v_0-v_1)\,(v_0-v_2)} + \frac{(x-v_0)\,(x-v_2)\,P_3(v_1)}{(v_1-v_0)\,(v_1-v_2)}$$
$$+\frac{(x-v_0)\,(x-x_1)\,P_3(v_2)}{(v_2-v_0)\,(v_2-v_1)} + (x-v_0)\,(x-v_1)\,(x-v_2)\,a_3.$$

IZA(3,3) with $v_0 = 0$, that contains one free parameter a_3, was used in [5].

`IZA(5,3):`
$$P_5(x) = \frac{(x-v_1)\,(x-v_2)\,P_5(v_0)}{(v_0-v_1)\,(v_0-v_2)} + \frac{(x-v_0)\,(x-v_2)\,P_5(v_1)}{(v_1-v_0)\,(v_1-v_2)} +$$
$$+\frac{(x-v_0)\,(x-v_1)\,P_5(v_2)}{(v_2-v_0)\,(v_2-v_1)}$$
$$+(x-v_0)\,(x-v_1)\,(x-v_2)\,\Big(a_3 + a_4 \overbrace{(x+v_0+v_1+v_2)}^{S_3^1} +$$
$$+a_5 \underbrace{(x^2+(v_0+v_1+v_2)\,x+{v_0}^2+v_1\,(v_0+v_1)+v_2(v_0+v_1+v_2))}_{S_3^2}\Big).$$

We mention that if $P_{p,r}(x)$, $p \geq r$, denotes the IZA(p,r) polynomial corresponding to r points

$$[v_0, y_0], [v_1, y_1] \ldots, [v_{r-1}, y_{r-1}]$$

with different abscissas, then one can easily verify that

$$P_{p,r}(x) = P_{p-1,r}(x) + Z_r(x) a_p S_r^{p-r}.$$

Hence for example if

$$P_{3,2}(x) = y_0 \frac{x-v_1}{v_0-v_1} + y_1 \frac{x-v_0}{v_1-v_0} + (x-v_0)(x-v_1)(a_2 + a_3(x+v_0+v_1))$$

we get

$$P_{4,2}(x) = P_{3,2}(x) + (x-v_0)(x-v_1) a_4 S_2^2.$$

Models based on the IZA representation can be used in various scenarios. One is the case, when the ordinates of the reference points are given. This scenario is common, and occurs e. g. when we want the approximant to go through some a priori given points. To the case, when the reference points are unknown and all of them are used to ensure the smooth transition, is devoted the next section. However, before moving to it we derive further forms for IZA polynomials based on the next lemma that considers the I and ZA components of the representation.

Lemma 3.3.
a) Let $p \geq r$, $\boldsymbol{\alpha} = (a_r, a_{r+1}, \ldots, a_p)^T$ and $\mathbf{w}_r(x) = (w_{0,r}(x), \ldots, w_{p-r,r}(x))^T$, where $w_{j,r}(x) = Z_r(x) S_r^j$ and S_r^j is defined by (11), $j = \overline{0, p-r}$. Then

$$Z_r(x) A_{p-r,r}(x) = \mathbf{w}_r(x)^T \cdot \boldsymbol{\alpha}. \tag{15}$$

b) Let $\mathbf{J}(x) = \begin{pmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{pmatrix}$ be a vector with $p+1$ elements, where $\mathbf{J}_1 = (1, x, \ldots, x^{r-1})^T$, $\mathbf{J}_2 = (x^r - w_{0,r}(x), x^{r+1} - w_{1,r}(x), \ldots, x^p - w_{p-r,r}(x))^T$ and $\mathbf{a} = (a_0, a_1, \cdots, a_p)^T$. Then

$$I_{r-1}(x) = \mathbf{J}(x)^T \cdot \mathbf{a}. \tag{16}$$

Proof. b) We have to show that $I_{r-1}(x) = P_{\bar{\boldsymbol{\alpha}}}(x) + \mathbf{J}_2^T \cdot \boldsymbol{\alpha}$, where $\bar{\boldsymbol{\alpha}} = (a_0, a_1, \ldots, a_{r-1})^T$, since $P_{\bar{\boldsymbol{\alpha}}}(x) + \mathbf{J}_2^T \cdot \boldsymbol{\alpha} = \mathbf{J}(x)^T \cdot \mathbf{a}$. Subtracting $P_{\boldsymbol{\alpha}}(x)$ from both side of the IZA representation (13) we get

$$P_p(x) - P_{\boldsymbol{\alpha}}(x) = I_{r-1}(x) + Z_r(x) A_{p-r,r}(x) - P_{\boldsymbol{\alpha}}(x),$$

and it suffices to notice that $P_p(x) - P_{\boldsymbol{\alpha}}(x) = P_{\bar{\boldsymbol{\alpha}}}(x)$ and $Z_r(x) A_{p-r,r}(x) - P_{\boldsymbol{\alpha}}(x) = \mathbf{w}_r(x)^T \cdot \boldsymbol{\alpha} - P_{\boldsymbol{\alpha}}(x) = -\mathbf{J}_2^T \cdot \boldsymbol{\alpha}$. □

From the above lemma we get

Corollary 3.4. An IZA(p, r), $p \geq r$, polynomial $P_a(x)$ can be expressed in four equivalent ways

$$\begin{aligned} P_a(x) &= I_{r-1}(x) + Z_r(x)\ A_{p-r,r}(x), \\ &= I_{r-1}(x) + Z_r(x)\ \mathbf{S}^T \cdot \boldsymbol{\alpha}, \\ &= I_{r-1}(x) + \mathbf{w}_r(x)^T \cdot \boldsymbol{\alpha}, \qquad (17) \\ &= \mathbf{J}(x)^T \cdot \mathbf{a} + \mathbf{w}_r(x)^T \cdot \boldsymbol{\alpha} \qquad (18) \end{aligned}$$

where

$$I_{r-1}(x) = \boldsymbol{\lambda}(x)^T \cdot \boldsymbol{\phi},$$

$$\begin{aligned} \boldsymbol{\lambda}(x) &= [\Pi_0(x), \Pi_1(x), \ldots, \Pi_{r-1}(x)]^T, \\ \boldsymbol{\phi} &= [y_0, y_1, \ldots, y_{r-1}]^T \end{aligned}$$

and $\Pi_i(x)$ is defined by (4).

The first, base form is a functional expression without vectors and indicates the degree of the polynomial components. The second form reveals the structure of the representation thanks to $\mathbf{S}$. Our central model in the next section is set up based on (17) and $\mathbf{J}(x)$ is leveraged in deriving the statistical properties of the estimates of the model's parameters.

4. THE MODEL AND ITS STATISTICAL PROPERTIES

This section contains the main statistical results. It shows how to properly choose the reference points in the two-part model and proves the consistency and asymptotic normality of $\hat{\boldsymbol{\beta}}$ and its scaled version, respectively, in a successive estimation mode, when first is estimated $\boldsymbol{a}$ and then $\boldsymbol{\beta}$.

Consider the two-part model (2), $p,\, q \geq r-1 \geq 0$, with smooth conditions (1) and uncorrelated errors described at the beginning of the introduction. We recall that the principal part of (2) is the second model, where the eminent role is played by the r reference points of $\mathcal{R}$. They, their location and ordinates, are so far undefined. We show now how to select $\mathcal{R}$ to ensure the continuity and quasi smoothness between the two polynomials $P_{\mathbf{a}}(x)$ and $P_{\mathbf{b}}(x)$ in zero. Let

$$v_j = -j\tau \quad \text{and} \quad P_{\mathbf{a}}(v_j) = P_{\mathbf{b}}(v_j), \quad j = \overline{0, r-1},$$

where τ is a small positive real number. From the theory of numerical derivatives of functions, particularly polynomials, it is known that all of the derivatives to the order $r-1$ of $P_{\mathbf{a}}(x)$ can be computed in zero with accuracy of order τ based on

$$\mathcal{R}_{\mathbf{a}} = \{\, [v_j,\, P_{\mathbf{a}}(v_j)] : \ v_j = -j\tau,\ j = \overline{0, r-1}\, \}. \tag{19}$$

Due to $P_{\mathbf{a}}(v_j) = P_{\mathbf{b}}(v_j),\ j = \overline{0, r-1}$, the same can be said about $P_{\mathbf{b}}(x)$. Consequently the derivatives of $P_{\mathbf{a}}(x)$ and $P_{\mathbf{b}}(x)$ should equal with accuracy τ and the quasi-spline conditions from (1)

$$P_{\mathbf{a}}^{(j)}(0) = P_{\mathbf{b}}^{(j)}(0) + o(\tau), \quad j = \overline{1,\, r-1},\ \ r \geq 2,$$

hold.

After choosing the set of reference points as $\mathcal{R} \equiv \mathcal{R}_{\mathbf{a}}$, where $\mathcal{R}_{\mathbf{a}}$ is defined by (19), the first two components of the second model of (2)

$$\tilde{y}_{i,N} = I_{\mathcal{R}}(j/N) + \mathbf{w}_r^T(j/N) \cdot \boldsymbol{\beta} + \varepsilon_{j,N}, \qquad \overline{1,\, N}, \tag{20}$$

can be expressed definitely based on this set by

$$I_{\mathcal{R}_{\mathbf{a}}}(x) = \frac{(x-v_1)\dots(x-v_{r-1})}{(v_0-v_1)\dots(v_0-v_{r-1})} P_{\mathbf{a}}(v_0) + \dots + \frac{(x-v_0)\dots(x-v_{r-2})}{(v_{r-1}-v_0)\dots(v_{r-1}-v_{r-2})} P_{\mathbf{a}}(v_{r-1})$$

and

$$\mathbf{w}_r(x) = (w_{0,r}(x), \dots\ w_{q-r,r}(x))^T,$$

where $w_{j,r}(x) = Z_r(x) S_r^j$ is defined in lemma 3.3, $j = \overline{0, q-r}$ and $\boldsymbol{\beta} = (b_r, b_{r+1}, \dots, b_q)$. We can see that this model corresponds to the third form (17) of the IZA(q, r) representation of the polynomial $P_{\mathbf{b}}(x)$ with $\mathcal{R} \equiv \mathcal{R}_{\mathbf{a}}$.

The matrix form of the two-part model (2) with $\mathcal{R}_{\mathbf{a}}$ is

$$\begin{aligned} \tilde{\mathbf{Y}}^* &= \mathbf{X}^* \mathbf{a} + \varepsilon*, \\ \tilde{\mathbf{Y}} &= \mathbf{I} + \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \end{aligned} \tag{21}$$

where

$$\begin{aligned}\tilde{\mathbf{Y}}^* &= \tilde{\mathbf{Y}}^*_{M\times 1}, \quad \mathbf{X}^* = \mathbf{X}^*{}_{M\times(p+1)}, \quad \mathbf{a} = \mathbf{a}_{(p+1)\times 1}, \quad \boldsymbol{\varepsilon}^* = \boldsymbol{\varepsilon}^*{}_{M\times 1}, \\ \tilde{\mathbf{Y}} &= \tilde{\mathbf{Y}}_{N\times 1}, \quad \mathbf{I} = \mathbf{I}_{N\times 1} = \{\, I_{\mathcal{R}_{\mathbf{a}}}(j/N)\,\}_{(j=\overline{1,N})\times 1}, \\ \mathbf{W} &= \mathbf{W}_{N\times l} = \{\, \mathbf{w}_r^T(j/N)\,\}_{(j=\overline{1,N})\times l}, \quad \boldsymbol{\beta} = \boldsymbol{\beta}_{l\times 1}, \quad \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_{N\times 1},\end{aligned}$$

and $l = q - r + 1$.

The successive estimates of the vectors $\mathbf{a}$ and $\boldsymbol{\beta}$ can be constructed by various techniques however we will assume that they are given by

Assumptions 1. Let assess the parameter $\mathbf{a}$ from the first model of (21) by the classic LS estimate

$$\widehat{\mathbf{a}} = (\mathbf{X}^{*\mathbf{T}}\mathbf{X}^*)^{-\mathbf{1}}\mathbf{X}^{*\mathbf{T}}\tilde{\mathbf{Y}}^*,$$

and $\boldsymbol{\beta}$ from the second model by the IZA LS esimate

$$\widehat{\boldsymbol{\beta}}_N = (\mathbf{W}^{\mathbf{T}}\mathbf{W})^{-\mathbf{1}}\mathbf{W}^{\mathbf{T}}(\tilde{\mathbf{Y}} - \hat{\mathbf{I}}\,), \tag{22}$$

where $\hat{\mathbf{I}} = \hat{\mathbf{I}}_{N\times 1} = \{\, I_{\mathcal{R}_{\hat{\mathbf{a}}}}(j/N)\,\}_{(j=\overline{1,N})\times 1}$ and

$$\mathcal{R}_{\hat{\mathbf{a}}} = \{\, [v_j,\, P_{\hat{\mathbf{a}}}(v_j)] : \;\; v_j = -j\tau, \;\; j = \overline{0, r-1}\,\}. \tag{23}$$

Before proving the consistency and asymptotic normality of the IZA LS estimate $\widehat{\boldsymbol{\beta}}_N$, derived from the second part of (21) that corresponds to the model (20) with an incomplete interpolation term, let us make some remarks. The Waring–Euler–Lagrange interpolating polynomial has a bad reputation because of its computational complexity and the Runge phenomenon in the case of equispaced nodes. However as it is shown in [18] these disadvantages can be handled using the barycentric interpolation formula and Tchebyshev nodes. Further, since the computational complexity of the LS normal system increases with growing polynomial degree, one should find a compromise for the values of r and q that influence the smoothness and accuracy of the resulting approximant, respectively.

While the parameter $\mathbf{a}$ of the left polynomial $P_{\mathbf{a}}(x)$ is estimated independently of the measurements from the right interval $[0, 1]$, the approximant $\hat{P}_{\mathbf{b}}(x) = I_{\mathcal{R}_{\hat{\mathbf{a}}}}(x) + \boldsymbol{w}_r^T(x)\cdot\hat{\boldsymbol{\beta}}$ incorporates measurements from both intervals. We emphasize that the use of model (20), where the reference points (19) and consequently (23) play a key role, has several benefits:

- it introduces a reparameterization of the classical polynomial regression: the r reference points and the $q-r+1$ leading coefficients $\boldsymbol{\beta}$ of $\mathbf{b}$ make up the parameters of the model

- to compute the approximant of $P_{\mathbf{b}}(x)$ in successive mode it is sufficient to asses within the two-part model (2) $q - r + 1$ coefficients ($\widehat{\boldsymbol{\beta}}_N$) instead of $q + 1$ ($\widehat{\mathbf{b}}$)

- similarly to the Hermite splines it ensures that the two approximants behave themselves smoothly in their shared point without solving the well known spline equations.

Now let us see the base properties of $\hat{\boldsymbol{\beta}}_N$.

Lemma 4.1. Under the assumptions 1

a) $\hat{\boldsymbol{\beta}}_N$ is unbiased

$$E\,\hat{\boldsymbol{\beta}}_N = \boldsymbol{\beta},$$

b)

$$\begin{aligned} cov\,\hat{\boldsymbol{\beta}}_N &= \sigma_*^2\,(\mathbf{W}^T\mathbf{W})^{-1}\,\mathbf{O}\,(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\,\mathbf{O}^T\,(\mathbf{W}^T\mathbf{W})^{-1} \\ &+ \sigma^2(\mathbf{W}^T\mathbf{W})^{-1}, \end{aligned}$$

where

$$\mathbf{O} = \mathbf{O}_{l\times(p+1)} = \mathbf{W}^T\mathbf{J},$$

and $\mathbf{J} = \mathbf{J}_{N\times(p+1)} = \{\,\mathbf{J}^T(j/N)^T\,\}_{(j=\overline{1,N})\times(p+1)}$.

P r o o f. a) Subtracting $\hat{\mathbf{I}}$ from both side of (21) we get

$$\tilde{\mathbf{Y}} - \hat{\mathbf{I}} = \mathbf{I} - \hat{\mathbf{I}} + \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{J}(\mathbf{a} - \hat{\mathbf{a}}) + \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{24}$$

Hence $E(\tilde{\mathbf{Y}} - \hat{\mathbf{I}}) = \mathbf{W}\boldsymbol{\beta}$ and from (22) we get that $\hat{\boldsymbol{\beta}}_N$ is unbiased.

b) From (22) and (24) we obtain

$$\begin{aligned} \hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta} &= (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{J}(\mathbf{a} - \hat{\mathbf{a}}) + (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\boldsymbol{\varepsilon} \\ &= (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{O}(\mathbf{X}^{*\mathbf{T}}\mathbf{X}^*)^{-\mathbf{1}}\mathbf{X}^{*T}\boldsymbol{\varepsilon}^* + (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\boldsymbol{\varepsilon}. \end{aligned} \tag{25}$$

□

Thereout the assertion follows immediately since $\boldsymbol{\varepsilon}^*$ and $\boldsymbol{\varepsilon}$ are uncorrelated.

Lemma 4.2. The elements of $\mathbf{W}^T\mathbf{W}$, $\mathbf{W}^T\mathbf{J}$ and $\mathbf{X}^{*T}\mathbf{X}^*$ have order N, i. e.

$$\begin{aligned} \mathbf{W}^T\mathbf{W} &= \{O_{i,j}(N)\}_{i=\overline{1,l};\,j=\overline{1,l}}, \\ \mathbf{W}^T\mathbf{J} &= \{O_{i,j}(N)\}_{i=\overline{1,l};\,j=\overline{1,p+1}}, \\ \mathbf{X}^{*T}\mathbf{X}^* &= \{O_{i,j}(N)\}_{i=\overline{1,l};\,j=\overline{p+1,p+1}}. \end{aligned}$$

P r o o f. The basis functions $w_{j,r}(x) = Z_r(x)S_r^j,\ j = \overline{0, q-r}$ are polynomials $Q_{r+j}^{j+1}(x)$ of degree $r+j$ with finite coefficients. Therefore the kth raw of $\mathbf{W}$, $k = \overline{1,N}$, equals

$$Q_r^1\Big(\frac{k}{N}\Big),\ Q_{r+1}^2\Big(\frac{k}{N}\Big)\ \ldots,Q_q^l\Big(\frac{k}{N}\Big)$$

and so $\mathbf{W}^T\mathbf{W} = \{\sum_{k=1}^{N} Q^i_{r+i-1}(\frac{k}{N})\, Q^j_{r+j-1}(\frac{k}{N})\}_{i=\overline{1,l};\, j=\overline{1,l}}$. Taking into account that for any finite integer $m \geq 1$

$$\sum_{k=1}^{N} \left(\frac{k}{N}\right)^m = O(N),$$

we get

$$\sum_{k=1}^{N} Q^i_{r+i-1}(\frac{k}{N})\, Q^j_{r+j-1}(\frac{k}{N}) = \sum_{k=1}^{N} \left(c_0 + \ldots + c_{2r+i+j-2} \left(\frac{k}{N}\right)^{2r+i+j-2} \right) = O(N),$$

and so $\mathbf{W}^T\mathbf{W} = \{O_{i,j}(N)\}_{i=\overline{1,l};\, j=\overline{1,l}}$.

The kth raw of $\mathbf{J}$, $k = \overline{1,N}$, equals

$$1, \left(\frac{k}{N}\right), \ldots \left(\frac{k}{N}\right)^{r-1}, \left(\frac{k}{N}\right)^{r} - w_{0,r}(\frac{k}{N}), \ \ldots, \ \left(\frac{k}{N}\right)^{p} - w_{p-r,r}(\frac{k}{N}),$$

and through the same argumentation as above we get $\mathbf{W}^T\mathbf{J} = \{O_{i,j}(N)\}_{i=\overline{1,l};\, j=\overline{1,p+1}}$. Similarly $\mathbf{X}^{*T}\mathbf{X}^* = \{O_{i,j}(N)\}_{i=\overline{1,l};\, j=\overline{p+1,p+1}}$. □

Theorem 4.3. Under the assumptions 1 the estimate $\hat{\boldsymbol{\beta}}_N$ is consistent,

$$\hat{\boldsymbol{\beta}}_N \ \xrightarrow{P} \ \boldsymbol{\beta} \quad \text{as} \quad N \to \infty.$$

P r o o f. We know from lemma 4.2 that the elements of $\mathbf{W}^T\mathbf{W}$ have order N. So the determinant of $\mathbf{W}^T\mathbf{W}$ and its subdeterminants have order of N^{p-r+1} and N^{p-r}, respectively. Consequently $(\mathbf{W}^T\mathbf{W})^{-1} = \{O_{i,j}(1/N)\}_{i=\overline{1,l};\, j=\overline{1,l}}$. Similarly $(\mathbf{X}^{*T}\mathbf{X}^*)^{-1} = \{O_{i,j}(1/N)\}_{i=\overline{1,l};\, j=\overline{p+1,p+1}}$. Hence and from lemmas 4.1, 4.2 we conclude that

$$\begin{aligned} cov\,\hat{\boldsymbol{\beta}}_N &= \sigma_*^2\,(\mathbf{W}^T\mathbf{W})^{-1}\,\mathbf{O}\,(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\,\mathbf{O}^T\,(\mathbf{W}^T\mathbf{W})^{-1} \\ &+ \sigma^2(\mathbf{W}^T\mathbf{W})^{-1} = \{O_{i,j}(1/N)\}_{i=\overline{1,l};\, j=\overline{p+1,p+1}}. \end{aligned}$$

Therefore from the Tchebychev inequality we get for any element $\hat{\boldsymbol{\beta}}_{i,N}$, $i = \overline{1, p-r+1}$ of $\hat{\boldsymbol{\beta}}_N$ and $\forall \epsilon \geq 0$

$$\lim_{N\to\infty} \mathbf{P}\{|\widehat{\boldsymbol{\beta}}_{i,N} - \boldsymbol{\beta}_i| > \epsilon\} \leq \lim_{N\to\infty} \frac{var\,\widehat{\boldsymbol{\beta}}_{i,N}}{\epsilon^2} = 0.$$

□

Notations 1. Let for $N \to \infty$

$$\frac{\mathbf{X}^{*T}\mathbf{X}^*}{N} \to \kappa\mathbf{U}, \quad \frac{\mathbf{W}^T\mathbf{J}}{N} \to \mathbf{c}, \quad \frac{\mathbf{W}^T\mathbf{W}}{N} \to \mathbf{V},$$

where $\mathbf{U} = \mathbf{U}_{(p+1)\times(p+1)}$, $\mathbf{c} = \mathbf{c}_{(q-r+1)\times(q-r+1)}$ and $\mathbf{V} = \mathbf{V}_{(q-r+1)\times(q-r+1)}$ are positive definite.

Since $\mathbf{X}^*$ and $\mathbf{W}$ are constructed based on the equispaced grids

$$\Big\{x_{i,M} = -\frac{i}{M},\ i=\overline{1,M}\Big\} \quad \text{and} \quad \Big\{x_{j,N} = \frac{j}{N},\ j=\overline{1,N}\Big\},$$

respectively, hence and from lemma 4.2 one can conclude that the above limits exist.

Within the framework of the two-part IZA scheme with quasi smooth transition the normality of a recursive estimate of a cubic, three reference points model's parameter was first proved in [13]. The next theorem generalizes that result for non recursive multidimensional estimators.

Theorem 4.4. Under assumptions 1 and notations 1

$$\Big(\sqrt{N}(\hat{\boldsymbol\beta}_N - \boldsymbol\beta)\Big) \ \xrightarrow{D}\ \mathcal{N}(\mathbf{0},\ \frac{\sigma_*^2}{\kappa}\,\mathbf{V}^{-1}\,\mathbf{c}\,\mathbf{U}^{-1}\,\mathbf{c}^T\,\mathbf{V}^{-1} + \sigma^2\,\mathbf{V}^{-1})$$

when $N \to \infty$.

Proof. Let us denote the i^{th} raw of $\mathbf{G}_N = (\frac{\mathbf{W}^T\mathbf{W}}{N})^{-1}$ by $\mathbf{g}_i^T$. Based on (25) the ith element of $N(\hat{\boldsymbol\beta} - \boldsymbol\beta)$,

$$\mathbf{g}_i^T\mathbf{O}(\mathbf{X}^{*\mathbf{T}}\mathbf{X}^*)^{-\mathbf{1}}\mathbf{X}^{*T}\boldsymbol\varepsilon^* + \mathbf{g}_i^T\mathbf{W}^T\boldsymbol\varepsilon,$$

is a sum of $K = M+N = (\kappa+1)N$ independent random variables. Let us denote it by

$$S_K = \xi_{1,K} + \ldots + \xi_{K,K},$$

where

$$\xi_{j,K} = \begin{cases} [\mathbf{g}_i^T\mathbf{O}(\mathbf{X}^{*\mathbf{T}}\mathbf{X}^*)^{-\mathbf{1}}\mathbf{X}^{*T}]_j\,\varepsilon_j^*\,, & 1 \le j \le M, \\ [\mathbf{g}_i^T\mathbf{W}^T]_{j-M}\,\varepsilon_{j-M}\,, & M+1 \le j \le N, \end{cases}$$

and $[\cdot]_J$, $J = j,\ j-M$, indicates the Jth column of the given matrix, variance of which is $O(K)$.

To prove that

$$e_K \equiv \frac{S_K}{\sqrt{var(S_K)}} \ \xrightarrow{D}\ \mathcal{N}(0,1) \quad \text{as } K \to \infty,$$

it suffices to show that the Lindeberg condition

$$\lim_{K\to\infty} \frac{1}{var S_K} \sum_{k=1}^{K} \int\limits_{\{x:\,|x-E\xi_{k,K}|\ge\epsilon\sqrt{var S_K}\}} (x - E\xi_{k,K})^2\,\mathrm{d}F_{k,K}(x) = 0 \tag{26}$$

holds. Since $E\xi_{k,K} = 0$ and $var\,S_K = O(K)$ we obtain

$$\frac{1}{var S_K}\sum_{k=1}^{K}\int\limits_{\{x:\,|x|\ge\epsilon\sqrt{var S_K}\}} x^2\,\mathrm{d}F_{k,K}(x) \le \frac{K}{O(K)}\max_{1\le k\le K}\int\limits_{\{x:\,|x|\ge\epsilon O(\sqrt{K})\}} x^2\,\mathrm{d}F_{k,K}(x).$$

Because $var\xi_{k,K}$ is finite and $\{x:\ |x| \geq \epsilon O(\sqrt{K})\} \downarrow \emptyset$ for $K \to \infty$, we have

$$\lim_{K\to\infty} \int\limits_{\{x:\,|x|\geq\epsilon O(\sqrt{K})\}} x^2\,\mathrm{d}F_{k,K}(x) = 0,$$

and the condition (26) is satisfied. Hence and from observation that

$$var\left(\sqrt{N}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})\right) \ \to\ \frac{\sigma_*^2}{\kappa}\,\mathbf{V}^{-1}\,\mathbf{c}\,\mathbf{U}^{-1}\,\mathbf{c}^T\,\mathbf{V}^{-1} + \sigma^2\,\mathbf{V}^{-1}$$

and

$$\sqrt{N}(\hat{\beta}_{N,i} - \beta_i) = \frac{S_K}{\sqrt{N}} = e_K\,\frac{\sqrt{varS_K}}{\sqrt{N}},$$

the proof is completed by standard techniques. □

We finish the section with results about the estimator

$$\hat{\mathbf{Y}} = \hat{\mathbf{I}} + \mathbf{W}\hat{\boldsymbol{\beta}} \tag{27}$$

$$= \hat{\mathbf{I}} + \mathbf{W}(\mathbf{W}^T\mathbf{W})^{-\mathbf{1}}\mathbf{W}^T(\tilde{\mathbf{Y}} - \hat{\mathbf{I}}). \tag{28}$$

Lemma 4.5. a) $\hat{\mathbf{Y}}$ is unbiased

$$E\hat{\mathbf{Y}} = \mathbf{Y},$$

b) Let $\mathcal{I}$ be the identical matrix and $\mathbf{H} \equiv \mathbf{H}_{N\times N} = \mathbf{W}(\mathbf{W}^T\mathbf{W})^{-\mathbf{1}}\mathbf{W}^T$.
Then

$$cov\,\hat{\mathbf{Y}} = \sigma_*^{\mathbf{2}}(\mathcal{I} - \mathbf{H})\mathbf{J}(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{J}^T(\mathcal{I} - \mathbf{H}) + \sigma^2\mathbf{H}.$$

c) The mean of the residual sum-of-squares is

$$E\ (\tilde{\mathbf{Y}} - \hat{\mathbf{Y}})^T \cdot (\tilde{\mathbf{Y}} - \hat{\mathbf{Y}}) = \sigma_*^2\ tr\left((\mathcal{I} - \mathbf{H})\mathbf{J}(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{J}^T\right) + \sigma^2(N - q + r - 1).$$

Proof. a) From (27) we get $E\hat{\mathbf{Y}} = \mathbf{Ja} + \mathbf{W}\boldsymbol{\beta} = \mathbf{Y}$.

b) Based on (28) and (24) we obtain

$$\begin{aligned}\hat{\mathbf{Y}} &= \hat{\mathbf{I}} + \mathbf{H}(\tilde{\mathbf{Y}} - \hat{\mathbf{I}})\\ &= \mathbf{J}\hat{\mathbf{a}} + \mathbf{H}(\mathbf{J}(\mathbf{a} - \hat{\mathbf{a}}) + \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon})\\ &= (\mathcal{I} - \mathbf{H})\mathbf{J}\hat{\mathbf{a}} + \mathbf{HY} + \mathbf{H}\boldsymbol{\varepsilon}.\end{aligned}$$

Hence the proof is completed immediately due to the fact that $\mathbf{H}$ is symmetric and idempotent, vectors $\hat{\boldsymbol{a}}, \boldsymbol{\varepsilon}$ are uncorrelated and $cov\,\hat{\boldsymbol{a}} = (\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\sigma_*^2$.

c) From (28) and $\mathbf{HW} = \mathbf{W}$ we get for the vector of residuals

$$\begin{aligned}\tilde{\mathbf{Y}} - \hat{\mathbf{Y}} &= \tilde{\mathbf{Y}} - \hat{\mathbf{I}} - \mathbf{H}\tilde{\mathbf{Y}} + \mathbf{H}\hat{\mathbf{I}} = (\mathcal{I} - \mathbf{H})\tilde{\mathbf{Y}} - (\mathcal{I} - \mathbf{H})\hat{\mathbf{I}}\\ &= (\mathcal{I} - \mathbf{H})(\tilde{\mathbf{Y}} - \mathbf{J}\hat{\mathbf{a}}) = (\mathcal{I} - \mathbf{H})\left(\mathbf{J}(\mathbf{a} - \hat{\mathbf{a}}) + \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\right)\\ &= (\mathcal{I} - \mathbf{H})\left(\mathbf{J}(\mathbf{a} - \hat{\mathbf{a}}) + \boldsymbol{\varepsilon}\right).\end{aligned}$$

Since

$$tr\,\mathbf{H} = tr\,\left(\mathbf{W^T W}(\mathbf{W^T W})^{-1}\right) = tr\,\mathcal{I}_{(q-r+1)\times(q-r+1)} = q - r + 1,$$

we have

$$tr\,(\mathcal{I} - \mathbf{H}) = N - q + r - 1.$$

Therefore

$$\begin{aligned} & E\ (\tilde{\mathbf{Y}} - \hat{\mathbf{Y}})^T \cdot (\tilde{\mathbf{Y}} - \hat{\mathbf{Y}}) = E\ tr\left[(\tilde{\mathbf{Y}} - \hat{\mathbf{Y}})(\tilde{\mathbf{Y}} - \hat{\mathbf{Y}})^T\right] \\ = \ & E\ tr\left[(\mathcal{I} - \mathbf{H})\left(\mathbf{J}(\mathbf{a} - \hat{\mathbf{a}}) + \varepsilon\right)\left(\mathbf{J}(\mathbf{a} - \hat{\mathbf{a}}) + \varepsilon\right)^T\right] \\ = \ & tr\left[(\mathcal{I} - \mathbf{H})(\ \mathbf{J}\, E\,(\mathbf{a} - \hat{\mathbf{a}})(\mathbf{a} - \hat{\mathbf{a}})^T \mathbf{J}^T + E\varepsilon\varepsilon^T\ )\right] \\ = \ & \sigma_*^2\ tr\left((\mathcal{I} - \mathbf{H})\mathbf{J}(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{J}^T\right) + \sigma^2(N - q + r - 1). \end{aligned}$$

□

5. SUMMARY

The paper introduced two novel concepts, the r-point transformation, and based on a special polynomial representation formula a two-part local regression model that provides approximants for neighboring polynomials with quasi smooth transition. We proved that the estimate of the model's main parameter and its scaled version is consistent and asymptotically normal, respectively. These properties confirm the validity of the proposed reference points based quasi smooth connection of two segments and the two-part approximation model.

While the r-point transformation is the tool that enabled a simple derivation of the representation, the corresponding regression model seems to be the main statistical outcome, based on which new approximation schemes and tools can be developed. The novel approach is not limited to two polynomials and 2D data. It can be generalized to global piecewise regression models for smoothing both 2D and 3D data with complex dependence.

ACKNOWLEDGEMENT

This work was partially supported by the European Regional Development Fund OP Research and Development Project *CaKS - Centre of excellence of Information sciences, ITMS Project Code: 26220120007*, the Slovak Research and Development Agency under contract *APVV-0035-10 Algorithms, Automata, and Discrete Data Structures* and the *VEGA project 1/0832/12*.

(Received April 30, 2012)

REFERENCES

[1] S. Csörgő and J. Mielniczuk: Nonparametric regression under long-range dependent normal errors. Ann. Statist. *23* (1995), 3, 1000–1014.

[2] N. D. Dikoussar: Function parametrization by using 4-point transforms. Comput. Phys. Commun. *99* (1997), 235–254.

[3] N. D. Dikoussar: Kusochno-kubicheskoje priblizhenije I sglazhivanije krivich v rezhime adaptacii. Communication JINR, P10-99-168, Dubna 1999.

[4] N. D. Dikusar: The basic element method. Math. Models Comput. Simulat. *3* (2011), 4, 492–507.

[5] N. D. Dikoussar and Cs. Török: Automatic knot finding for piecewise-cubic approximation. Matem. Mod. *18* (2006), 3, 23–40.

[6] N. D. Dikoussar and Cs. Török: On one approach to local surface smoothing. Kybernetika *43* (2007), 4, 533–546.

[7] R. L. Eubank: Nonparametric Regression and Spline Smoothing. Marcel Dekker, Inc., 1999.

[8] D. Kahaner, C. Moler, and S. Nash: Numerical Methods and Software. Prentice-Hall, Inc., 1989.

[9] T. Kepič, Cs. Török, and N. D. Dikoussar: Wavelet compression. In: 13. International Workshop on Computational Statistics, Bratislava 2004, pp. 49–52.

[10] A. Matejčiková and Cs. Török: Noise suppression in RDPT. Forum Statisticum Slovacum *3* (2005), 199–203.

[11] E. A. Nadaraya: On estimating regression. Theory Probab. Appl. *9* (1964), 141–142.

[12] M. Révayová and Cs. Török: Piecewise approximation and neural networks. Kybernetika *43* (2007), 4, 547–559.

[13] M. Révayová and Cs. Török: Reference points based recursive approximation. Kybernetika *49* (2013), 1, 60–72.

[14] B. D. Riplay: Pattern Recognition and Neural Networks. Cambridge University Press 1996.

[15] G. A. F. Seber: Linear Regression Analysis. J. Wiley and Sons, New York 1977.

[16] Cs. Török: 4-point transforms and approximation. Comput. Phys. Commun. *125* (2000), 154–166.

[17] Cs. Török and N. D. Dikoussar: Approximation with DPT. Comput. Math. Appl. *38* (1999), 211–220.

[18] L. N. Trefethen: Approximation Theory and Approximation Practice. SIAM, 2013.

Csaba Török, Pavol Jozef Šafárik University in Košice, Jesenná 5, 040 01 Košice. Slovak Republic.

e-mail: csaba.torok@tuke.sk