KYBERNETIKA — VOLUME *49* (2013), NUMBER 1, PAGES 60–72

REFERENCE POINTS BASED RECURSIVE APPROXIMATION

Martina Révayová and Csaba Török

The paper studies polynomial approximation models with a new type of constraints that enable to get estimates with significant properties. Recently we enhanced a representation of polynomials based on three reference points. Here we propose a two-part cubic smoothing scheme that leverages this representation. The presence of these points in the model has several consequences. The most important one is the fact that by appropriate location of the reference points the resulting approximant of two successively assessed neighboring approximants will be smooth. We also show that the considered models provide estimates with appropriate statistical properties such as consistency and asymptotic normality.

Keywords: approximation model, consistency, asymptotic normality

Classification: 41A10, 62-07, 62F10, 62J05, 62L12, 65D05, 65D07, 65D10

1. INTRODUCTION

Data with complex, nonlinear structure can be approximated using various nonparametric methods such as kernel smoothers [6] or smoothing splines [7]. There are situations when for further work with data or the applied model one needs approximation equations that these methods do not provide. The paper proposes smoothing schemas with a new parametric approximation model for noisy data in plane, where the parameter is estimated by a Robbins–Monro type recursive sequence [11]. Recursive estimating procedures occur e. g. in stochastic approximation [11] or in adaptive filters, see the recursive least squares filter [5]. These methods are utilized in practice for solving problems, where the information extraction and noise reduction proceed online, and use the accumulated and most recently arrived data effectively. We were motivated by the knot detection problem for piecewise approximation, when the detection algorithm proceeds successively from left to right [3]. After finding a knot, one can proceed at the first glance independently of the past, however it is advisable to take the current state into consideration, too. In our case it means smooth transition from the last segment to the new one. Given the model and the recursive estimate, our goal is to investigate the properties of this sequence.

In the background of the proposed smoothing schemas lies the discrete projective transformation, introduced in [1] and studied in [4, 10]. Using the complete version of this transformation we succeeded to derive in [8] an enhanced representation of polynomials using three reference points. This representation enables elaborating new approximating polynomial models, in which the first three coefficients a_0, a_1, a_2 of a polynomial $P_p(x) = a_0 + a_1 x + \cdots + a_p x^p$, $p \geq 3$, are replaced with new parameters defined by the reference points. Thanks to the reference points the resulting models possess in addition to approximating properties also interpolating ones: the three new parameters are used for interpolation and the rest for the approximation. For the free parameter of these models we construct recursive schemas that are similar to the ones in [2, 3]. Our goal was to improve those approximating procedures and investigate their theoretical properties. The main outcomes of this paper are the quasi smooth approximating two-part schema and the proof of the asymptotic normality of the recursive estimate.

The structure of the paper is the following. In section two we provide the three reference point representation of polynomials. In the next two sections we introduce two cubic smoothing models for which we construct iterative estimations and prove their consistency and asymptotic normality. The main results are in section four. Section five contains a simulation study.

2. REPRESENTATION OF POLYNOMIALS

In our approach to smoothing noisy data by piecewise polynomial models with a special type of constraints the key role is played by a set of three arbitrary, but mutually different reference points

$$\mathcal{R} = \{[v_0, y_0], [v_1, y_1], [v_2, y_2] : \; v_0 \neq v_1 \neq v_2, \; v_i \in \mathbb{R}, i = 0, 1, 2\}$$

and a special representation of polynomials using these points. The aim of this section is to present this representation introduced in [8].

We will denote the polynomials

$$P_{\mathbf{a},p}(x) = \sum_{i=0}^{p} a_i x^i$$

for the sake of simplicity both with $P_{\mathbf{a}}(x)$ and $P_p(x)$, depending on whether we want to emphasize the coefficients $\mathbf{a}^T = (a_0, a_1, \ldots, a_p)$ or the polynomial degree p. The next theorem provides the $P = I + ZA$ representation of polynomials.

Theorem 2.1. Let p be an integer greater than or equal to three. Then the polynomial $P_p(x)$ can be expressed based on the set $\mathcal{R}$ with $y_i = P_p(v_i)$, $i = 0, 1, 2$, in the form

$$P_p(x) = I(x) + Z(x)A(x), \tag{1}$$

where

$$\begin{aligned} I(x) \equiv I^{\mathcal{R}}(x) &= p_0(x)y_0 + p_1(x)y_1 + p_2(x)y_2, & (2)\\ Z(x) \equiv Z^{\mathcal{R}}(x) &= (x - v_0)(x - v_1)(x - v_2), & (3)\\ A(x) &= \sum_{i=3}^{p} a_i T_{i-3}(x) \end{aligned}$$

and

$$p_0(x) = \tfrac{(x-v_1)(x-v_2)}{(v_0-v_1)(v_0-v_2)}, \; p_1(x) = \tfrac{(x-v_0)(x-v_2)}{(v_1-v_0)(v_1-v_2)}, \; p_2(x) = \tfrac{(x-v_0)(x-v_1)}{(v_2-v_0)(v_2-v_1)},$$

$$T_i(x) = \sum_{j=0}^{i} r_j x^{i-j},\ i \in Z_0^+,$$

$$r_j = \sum_{k_0=0}^{j} \sum_{k_1=0}^{j-k_0} v_0^{k_0} v_1^{k_1} v_2^{j-k_0-k_1},\ j \in Z_0^+.$$

Let us notice that I is an incomplete interpolating polynomial of degree 2, A is a polynomial of degree $p-3$, and

$$I(v_0) = y_0, I(v_1) = y_1, I(v_2) = y_2,$$

$$Z(v_0) = 0, Z(v_1) = 0 \text{ and } Z(v_2) = 0.$$

Based on this representation with mutually different reference points $[v_0, y_0]$, $[v_1, y_1]$ and $[v_2, y_2]$, where $y_i = P_3(v_i)$, $i = 0, 1, 2$, we will consider a new parametrization for a cubic polynomial

$$\begin{aligned} P_3(x) &= b_0 + b_1 x + b_2 x^2 + b_3 x^3 \\ &= I(x) + Z(x) b_3 \\ &= p_0(x) y_0 + p_1(x) y_1 + p_2(x) y_2 + Z(x) b_3, \end{aligned} \tag{4}$$

where from the original four parameters remains only the parameter b_3. We will take advantage of this representation of cubic polynomials (4) in the next sections in the definition of the approximating models.

3. MODEL WITH KNOWN REFERENCE POINTS

In this and the subsequent section we introduce two cubic smoothing schemas that are obtained from the polynomial representation (4). We will assume about the model of this section that the ordinates of the three reference points are known. This one-part model does not depend on the choice of the abscissas of the reference points. In the two-part model of the next section the three ordinates are unknown and they have to be assessed, and the abscissas are not arbitrary, but bounded to the shared point.

3.1. Model

Consider a cubic polynomial $P_{\mathbf{b}}(x)$, $x \in [0,\ 1]$, where $\mathbf{b}^T = (b_0, b_1, b_2, b_3)$ and a set of N data points in the interval $[0,\ 1]$

$$\left\{ [x_{i,N}, \widetilde{y}_{i,N}] : x_{i,N} = \frac{i}{N},\ \widetilde{y}_{i,N} \equiv P_{\mathbf{b}}(x_{i,N}) + \varepsilon_{i,N}, i = \overline{1, N} \right\},$$

where $\{\varepsilon_{i,N}\}_{i=\overline{1,N}}$ is an uncorrelated error sequence, $E\varepsilon_{i,N} = 0$ and $E\varepsilon_{i,N}\varepsilon_{j,N} = \sigma^2 \delta_{i,j} < \infty$ $i, j = \overline{1, N}$, $\delta_{i,j}$ is the Kronecker delta.

The definition of the new approximation model for these data needs three reference points. In this section we will suppose that they are given and errorless. Let the set of reference points be

$$\mathcal{R} = \{[v_0, y_0],\ [v_1, y_1],\ [v_2, y_2] : v_0 \neq v_1 \neq v_2,\ y_i = P_{\mathbf{b}}(v_i), i = 0, 1, 2\}. \tag{5}$$

Based on (4) we define a cubic, the simplest approximating model

$$\widetilde{y}(x) = I(x) + Z(x)b + \epsilon, \tag{6}$$

where $I(x)$, $Z(x)$ are defined by (2), (3) using v_i and y_i, $i = 0, 1, 2$, from (5) and $b \equiv b_3$.

3.2. Recursive estimate and its properties

For the approximation of $P_{\mathbf{b}}(x)$ according to (6) let us consider

$$\widehat{y}(x) = I(x) + Z(x)\widehat{b},$$

where for the evaluation of the estimate $\widehat{b}$ various techniques can be considered. We will use a recursive estimation scheme, see [2, 3],

$$b_i = b_{i-1} + \frac{Z_{i,N}}{\zeta_i}(\tilde{y}_{i,N} - I_{i,N} - Z_{i,N}b_{i-1}), \qquad i = \overline{1,N},\ b_0 = 0, \tag{7}$$

where $\tilde{y}_{i,N} = \tilde{y}(x_{i,N})$, $I_{i,N} = I(x_{i,N})$, $Z_{i,N} = Z(x_{i,N}) = \left(\frac{i}{N} - v_0\right)\left(\frac{i}{N} - v_1\right)\left(\frac{i}{N} - v_2\right)$ and $\zeta_i = \sum_{k=1}^{i} Z_{k,N}^2$. Using (6) b_i can be expressed as

$$b_i = b_{i-1} + \frac{Z_{i,N}}{\zeta_i}(Z_{i,N}(b - b_{i-1}) + \varepsilon_{i,N}),$$

whereas

$$b_1 = b + \frac{Z_{1,N}\varepsilon_{1,N}}{\zeta_1}, \quad b_2 = b + \frac{Z_{1,N}\varepsilon_{1,N} + Z_{2,N}\varepsilon_{2,N}}{\zeta_2}.$$

Hence by induction we get a relation for b_i

$$b_i = b + \frac{\sum\limits_{k=1}^{i} Z_{k,N}\varepsilon_{k,N}}{\zeta_i}, \qquad i = \overline{1,N}, \tag{8}$$

that enables the investigation of its properties.

Let us now take a look at the basic statistical properties of the estimates b_i and $\widehat{b} \equiv b_N$. From (8) it follows immediately that

$$Eb_i = b, \quad i = \overline{1,N}.$$

Mentioning that

$$\zeta_N = \frac{1}{N^6}\sum_{k=1}^{N}((k - Nv_0)(k - Nv_1)(k - Nv_2))^2 = O(N), \tag{9}$$

from (8) we also get

$$Db_N = \frac{\sigma^2\zeta_N}{\zeta_N^2} = \frac{1}{O(N)}.$$

Hence and from the Tchebychev inequality it follows that for $\forall\epsilon > 0$

$$\lim_{N\to\infty} \mathbf{P}\{|b_N - b| > \epsilon\} \le \lim_{N\to\infty} \frac{Db_N}{\epsilon^2} = 0,$$

i. e. the estimate b_N is consistent, $b_N \xrightarrow{P} b$. However we can prove the asymptotic normality, too.

Theorem 3.1. Consider the estimate (7) of the free parameter b of the model (6). Then

$$(b_N - b)\frac{\sqrt{\zeta_N}}{\sigma} \xrightarrow{d} \mathcal{N}(0,1),$$

where ζ_N is defined by (9).

P r o o f. First we show that

$$(b_N - b)\frac{\sqrt{\zeta_N}}{\sigma} = \frac{S_N}{\sqrt{DS_N}},$$

where

$$\xi_{k,N} = \frac{\sqrt{\zeta_N}}{\sigma\sqrt{N}} Z_{k,N}\varepsilon_{k,N}, \quad k = \overline{1,N},$$

$$S_N \;=\; \xi_{1,N} + \xi_{2,N} + \ldots + \xi_{N,N}.$$

Because $S_N = \frac{\sqrt{\zeta_N}}{\sigma\sqrt{N}} \sum_{k=1}^{N} Z_{k,N}\varepsilon_{k,N}$, we have $DS_N = \frac{1}{N}\zeta_N^2$. Hence and from (8) for $i = N$ it follows that

$$(b_N - b)\frac{\sqrt{\zeta_N}}{\sigma} = \frac{\sqrt{\zeta_N}\sum\limits_{k=1}^{N} Z_{k,N}\varepsilon_{k,N}}{\sigma\zeta_N} = \frac{S_N}{\sqrt{DS_N}}.$$

To prove that this estimate has a standard normal distribution for $N \to \infty$, it is sufficient to show that for $\forall\epsilon > 0$ the Lindeberg condition

$$\lim_{N\to\infty} \frac{1}{DS_N} \sum_{k=1}^{N} \int_{\{x:|x-E\xi_{k,N}|\ge\epsilon\sqrt{DS_N}\}} (x - E\xi_{k,N})^2 \,\mathrm{d}F_{k,N}(x) = 0 \tag{10}$$

is satisfied, where $F_{k,N}(x)$ is the distribution function of $\xi_{k,N}$. Since $E\xi_{k,N} = 0$ and (9) implies that $DS_N = O(N)$, we have

$$\frac{1}{DS_N}\sum_{k=1}^{N}\int_{\{x:|x|\ge\epsilon\sqrt{DS_N}\}} x^2\,\mathrm{d}F_{k,N}(x) \le \frac{N}{O(N)} \max_{1\le k\le N} \int_{\{x:|x|\ge\epsilon O(\sqrt{N})\}} x^2\,\mathrm{d}F_{k,N}(x).$$

For any k, $1 \le k \le N$, $\xi_{k,N}$ has a finite integral

$$D\xi_{k,N} = \int\limits_{-\infty}^{\infty} x^2\,\mathrm{d}F_{k,N}(x) < \infty.$$

It is left to prove that if the integrals are finite the remainings tend to zero. Obviously

$$\{x : |x| \geq \epsilon O(\sqrt{N})\} \downarrow \emptyset \text{ for } N \to \infty,$$

therefore

$$\lim_{N\to\infty} \int_{\{x:|x|\geq \epsilon O(\sqrt{N})\}} x^2 \, \mathrm{d}F_{k_N,N}(x) = 0,$$

and consequently the condition (10) is fulfilled. □

It was the case when the reference points were known. As we shall see in the next section, which deals with successive approximation over two intervals, the consistency and normality of an appropriately constructed estimate b_N hold even when the ordinates of the reference points are not given but assessed.

4. ESTIMATED REFERENCE POINTS

This section deals with the case when the ordinates of the reference points are unknown. We will investigate two polynomials over two neighboring intervals that are smooth in the shared point. First we describe a two-part scheme, and then construct an iterative estimate and investigate its properties.

4.1. Model

Consider two cubic polynomials $P_{\mathbf{a}}(x)$ and $P_{\mathbf{b}}(x)$ over the intervals $[-1, 0]$ and $[0, 1]$, respectively, where $\mathbf{a}^T = (a_0, a_1, a_2, a_3)$ and $\mathbf{b}^T = (b_0, b_1, b_2, b_3)$. Let be given two data sets from these intervals

$$\left\{[x_{i,M}, \widetilde{y}_{i,M}] : x_{i,M} = -\frac{i}{M}, \widetilde{y}_{i,M} \equiv P_{\mathbf{a}}(x_{i,M}) + \varepsilon^*_{i,M}, i = \overline{1, M}\right\},$$
$$\left\{[x_{i,N}, \widetilde{y}_{i,N}] : x_{i,N} = \frac{i}{N}, \widetilde{y}_{i,N} \equiv P_{\mathbf{b}}(x_{i,N}) + \varepsilon_{i,N}, i = \overline{1, N}\right\},$$

where $M, N \gg 1$. Let $M = \kappa N$, $0 < \kappa < \infty$ and $\{\varepsilon^*_{i,M}\}_{i=\overline{1,M}}$, $\{\varepsilon_{k,N}\}_{k=\overline{1,N}}$ be uncorrelated error sequences, $E\varepsilon^*_{i,M} = 0$, $E\varepsilon_{k,N} = 0$, $E\varepsilon^*_{i,M}\varepsilon^*_{j,M} = \sigma^2\delta_{i,j} < \infty$, $E\varepsilon_{k,N}\varepsilon_{l,N} = \sigma^2\delta_{k,l} < \infty$ and $E\varepsilon^*_{i,M}\varepsilon_{k,N} = 0$ for $i, j = \overline{1, M}$, $k, l = \overline{1, N}$, where $\delta_{*,*}$ is the Kronecker delta. We suppose that the polynomials are continuous and quasi smooth at $x = 0$

$$\begin{aligned} P_{\mathbf{b}}(0) &= P_{\mathbf{a}}(0), \\ P'_{\mathbf{b}}(0) &= P'_{\mathbf{a}}(0) + o(\tau), \\ P''_{\mathbf{b}}(0) &= P''_{\mathbf{a}}(0) + o(\tau), \end{aligned} \tag{11}$$

i. e. $|P^{(j)}_{\mathbf{a}}(0) - P^{(j)}_{\mathbf{b}}(0)| < c_j\tau$, where τ is a small positive real number and $P^{(j)}$ denotes the jth derivative of P. We propose a two-part smoothing scheme for data from the interval $[-1, 1]$. For the data points from $[-1, 0]$ we consider the classical polynomial model

$$\widetilde{y}(x) = P_{\mathbf{a}}(x) + \varepsilon^*. \tag{12}$$

Similarly to the preceding section we use the formula (4) for the definition of the local model

$$\widetilde{y}(x) = I^{\mathcal{R}}(x) + Z(x)b + \varepsilon \tag{13}$$

for the interval [0, 1], where $\mathcal{R}$ denotes a set of reference points. The abscissas of the reference points from $\mathcal{R}$ can be selected suitably to ensure the quasi continuity of transition of second order (11) in the shared zero point of the two local approximants. If we choose the abscissas of these points as $\{-2\tau, -\tau, 0\}$, where τ is a small positive real number, and suppose

$$P_{\mathbf{b}}(-2\tau) = P_{\mathbf{a}}(-2\tau), P_{\mathbf{b}}(-\tau) = P_{\mathbf{a}}(-\tau) \text{ and } P_{\mathbf{b}}(0) = P_{\mathbf{a}}(0), \tag{14}$$

then from the numerical mathematics it is well known that (11) holds. So let the set of reference points $\mathcal{R}$ be

$$\mathcal{R} = \{[-2\tau, P_{\mathbf{b}}(-2\tau)],\ [-\tau, P_{\mathbf{b}}(-\tau)],\ [0, P_{\mathbf{b}}(0)]\}.$$

Since we assume that (14) holds, we get

$$\mathcal{R} = \mathcal{R}_{\mathbf{a}} = \{[-2\tau, P_{\mathbf{a}}(-2\tau)],\ [-\tau, P_{\mathbf{a}}(-\tau)],\ [0, P_{\mathbf{a}}(0)]\}. \tag{15}$$

The models (12) and (13) with reference points from (15) make up our two-part smoothing scheme for $P_{\mathbf{a}}(x)$ and $P_{\mathbf{b}}(x)$ that satisfy the new type of constraints (11).

So we consider two polynomials over two neighboring intervals. We can estimate the coefficients of the polynomial from the first interval using any method, however for the investigation of the theoretical properties of the second model we will consider the LS method. As we shall see, the approximant for the second interval uses three estimated points from the approximant for the first interval as reference points.

4.2. Recursive estimate of b_i

This section deals with the approximation of $P_{\mathbf{b}}(x)$. Let us denote the LS approximant of $P_{\mathbf{a}}(x)$ from the interval [−1, 0] by $P_{\hat{\mathbf{a}}}(x)$. Based on (13) we introduce for $P_{\mathbf{b}}(x)$ an approximant

$$\hat{y}(x) = I^{\hat{\mathcal{R}}}(x) + Z(x)\hat{b}, \tag{16}$$

where $\hat{\mathcal{R}}$ denotes a set of reference points, the ordinates of which are estimated by $P_{\hat{\mathbf{a}}}(x)$:

$$\hat{\mathcal{R}} = \mathcal{R}_{\hat{\mathbf{a}}} = \{[-2\tau, P_{\hat{\mathbf{a}}}(-2\tau)],\ [-\tau, P_{\hat{\mathbf{a}}}(-\tau)],\ [0, P_{\hat{\mathbf{a}}}(0)]\}.$$

For the evaluation of $\hat{b}$ we propose the recursive scheme

$$b_i = b_{i-1} + \frac{Z_{i,N}}{\zeta_i}(\tilde{y}_{i,N} - \hat{I}_{i,N} - Z_{i,N}b_{i-1}),\ \ b_0 = 0,\ \ i = \overline{1, N}, \tag{17}$$

where based on (3)

$$Z_{i,N} \equiv Z^{\hat{\mathcal{R}}}(x_{i,N}) = \left(\frac{i}{N} + 2\tau\right)\left(\frac{i}{N} + \tau\right)\frac{i}{N},$$

$$\zeta_i = \sum_{k=1}^{i} Z_{k,N}^2, \quad \tilde{y}_{i,N} = \tilde{y}(x_{i,N}) \quad \text{and} \quad \hat{I}_{i,N} = I^{\hat{\mathcal{R}}}(x_{i,N}).$$

From (13) we get

$$b_i = b_{i-1} + \frac{Z_{i,N}}{\zeta_i}(I_{i,N} - \hat{I}_{i,N} + Z_{i,N}(b - b_{i-1}) + \varepsilon_{i,N}), \;\; b_0 = 0, \;\; i = \overline{1,N},$$

where $I_{i,N} \equiv I^{\mathcal{R}}_{i,N}$ and $\mathcal{R}$ is defined by (15). It can be shown analogously as (8) that

$$b_i = b + \frac{1}{\zeta_i}\left(\sum_{k=1}^{i} Z_{k,N}(I_{k,N} - \hat{I}_{k,N} + \varepsilon_{k,N})\right), i = \overline{1,N}. \tag{18}$$

While the recursive formula (17) serves for numerical computation of b_i, relation (18) for theoretical investigation of its properties.

We turn now to the exploration of the statistical properties of the estimate b_i. First we clear up the convergence rate of the sequence $\{b_i\}$ and then prove that the asymptotic distribution of its modification is the Gaussian one.

Based on (2) and (4) we express $I_{k,N} - \hat{I}_{k,N}$ the following way

$$\begin{aligned} I_{k,N} - \hat{I}_{k,N} &= \sum_{i=0}^{2} p_i(x_{k,N})(P_{\mathbf{a}}((i-2)\tau) - P_{\hat{\mathbf{a}}}((i-2)\tau) \\ &= (a_0 - \hat{a}_0) + (a_1 - \hat{a}_1)\frac{k}{N} + (a_2 - \hat{a}_2)\frac{k^2}{N^2} + (a_3 - \hat{a}_3)\left(-3\tau\frac{k^2}{N^2} - 2\tau^2\frac{k}{N}\right). \end{aligned}$$

Let us denote

$$\begin{aligned} \Delta\mathbf{a} &= \mathbf{a} - \hat{\mathbf{a}} = -(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\varepsilon^*, \qquad (19) \\ \boldsymbol{\nu}_N^T &= \left(\sum_{k=1}^{N} Z_{k,N}, \sum_{k=1}^{N}\frac{k}{N}Z_{k,N}, \sum_{k=1}^{N}\frac{k^2}{N^2}Z_{k,N}, \sum_{k=1}^{N}\left(-3\tau\frac{k^2}{N^2} - 2\tau^2\frac{k}{N}\right)Z_{k,N}\right) \\ &= \left(\frac{N}{4} + O(1), \frac{N}{5} + O(1), \frac{N}{6} + O(1), \frac{\tau(4\tau+5)N}{10} + O(1)\right). \qquad (20) \end{aligned}$$

Then

$$\sum_{k=1}^{N} Z_{k,N}\left(I_{k,N} - \hat{I}_{k,N}\right) = \boldsymbol{\nu}_N^T\Delta\mathbf{a}.$$

Hence and from (18) we get

$$b_N = b + \frac{1}{\zeta_N}\left(\boldsymbol{\nu}_N^T\Delta\mathbf{a} + \sum_{k=1}^{N} Z_{k,N}\varepsilon_{k,N}\right), \tag{21}$$

where

$$\zeta_N = \sum_{k=1}^{N} Z_{k,N}^2 = \frac{1}{N^6}\sum_{k=1}^{N}((k+2N\tau)(k+N\tau)k)^2$$
$$= \left(\frac{1}{7}+\tau+\frac{13}{5}\tau^2+3\tau^3+\frac{4}{3}\tau^4\right)N+O(1). \quad (22)$$

Lemma 4.1. Consider the estimate (17) of the free parameter b of the model (13). Then

1. $Eb_i = b$ for $i = \overline{1,N}$,
2. $Db_N = \frac{1}{O(N)}$.

Proof. a) Since $E\hat{\mathbf{a}} = \mathbf{a}$, we have $E\hat{I}_{k,N} = I_{k,N}$ and so the statement is a consequence of (18).
b) From (21) we obtain

$$Db_N = \frac{1}{\zeta_N^2}\left(D\boldsymbol{\nu}_N^T\Delta\mathbf{a} + D\sum_{k=1}^{N}\varepsilon_{k,N}Z_{k,N}\right).$$

Since

$$var\Delta\mathbf{a} = (\mathbf{X}^T\mathbf{X})^{-1}\sigma^2 = \left\{\frac{1}{O_{ij}(M)}\right\}_{i,j=\overline{0,3}},$$

we get from (20) due to $M = \kappa N$ that

$$D\boldsymbol{\nu}_N^T\Delta\mathbf{a} = \frac{O(N^2)}{O(M)} = O(N). \quad (23)$$

(22) implies

$$D\sum_{k=1}^{N} Z_{k,N}\varepsilon_{k,N} = \sigma^2\zeta_N = O(N).$$

Hence and from (23), (21) we get $Db_N = \frac{O(N)}{O(N^2)}$. □

Let us introduce the notations

$$\mathbf{U}_M = \frac{\mathbf{X}^T\mathbf{X}}{M}, \mathbf{U} = \lim_{M\to\infty}\mathbf{U}_M, \mathbf{c}_N = \frac{\boldsymbol{\nu}_N}{N} \text{ and } \mathbf{c} = \lim_{N\to\infty}\mathbf{c}_N.$$

Theorem 4.2. If the assumptions of lemma 4.1 are fulfilled, then

1. the estimate b_N is consistent, $b_N \xrightarrow{P} b$,

2. $(b_N - b)\frac{\sqrt{\zeta_N}}{\sigma} \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{\kappa\mathbf{T}}\mathbf{c}^T\mathbf{U}^{-1}\mathbf{c} + 1\right)$,

where $\mathbf{T} = \frac{1}{7} + \tau + \frac{13}{5}\tau^2 + 3\tau^3 + \frac{4}{3}\tau^4$.

P r o o f. a) It follows from lemma 4.1 and the Tchebychev inequality.
b) First we show that

$$(b_N - b)\frac{\zeta_N}{\sigma\sqrt{N\delta_N}} \xrightarrow{d} \mathcal{N}(0,1),$$

where $\delta_N = \frac{1}{\kappa}\mathbf{c}_N^T\mathbf{U}_M^{-1}\mathbf{c}_N + \frac{\zeta_N}{N}$.
Let S_N be (see (21))

$$S_N = (b_N - b)\zeta_N = \boldsymbol{\nu}_N^T\Delta\mathbf{a} + \sum_{k=1}^{N} Z_{k,N}\varepsilon_{k,N}.$$

Using (19) we obtain

$$\boldsymbol{\nu}_N^T\Delta\mathbf{a} = \sum_{k=1}^{M} d_i\varepsilon_{k,M}^*,$$

and so S_N is a sum of non-correlated random variables with finite variances. From the relations $D\sqrt{M}\Delta\mathbf{a} = \mathbf{U}_M^{-1}\sigma^2$, $\frac{N}{M} = \frac{1}{\kappa}$ and the non-correlation of ε^* and ε we get

$$DS_N = \frac{\sigma^2}{M}\boldsymbol{\nu}_N^T\mathbf{U}_M^{-1}\boldsymbol{\nu}_N + \sigma^2\zeta_N = \sigma^2 N\left(\frac{1}{\kappa}\frac{\boldsymbol{\nu}_N^T}{N}\mathbf{U}_M^{-1}\frac{\boldsymbol{\nu}_N}{N} + \frac{\zeta_N}{N}\right) = \sigma^2 N\delta_N,$$

and so

$$(b_N - b)\frac{\zeta_N}{\sigma\sqrt{N\delta_N}} = \frac{S_N}{\sqrt{DS_N}}.$$

Hence thanks to $DS_N = O(N)$ the proof of the fulfillment of the Lindeberg condition is similar to the proof of theorem 3.1.

Since $\lim_{N\to\infty}\frac{N\delta_N}{\zeta_N} = \frac{1}{\kappa\mathbf{T}}\mathbf{c}^T\mathbf{U}^{-1}\mathbf{c} + 1$ we get by standard techniques

$$(b_N - b)\frac{\sqrt{\zeta_N}}{\sigma} \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{\kappa\mathbf{T}}\mathbf{c}^T\mathbf{U}^{-1}\mathbf{c} + 1\right). \qquad \square$$

We proved the consistency and asymptotic normality of the estimate (21) of the free parameter b and its modification in the local cubic model (13), respectively. Obviously based on the representation (1) we could construct polynomial models of higher degree, or tending τ to zero express $I(x)$ with the first two derivatives of $P_p(x)$.

5. SIMULATION STUDY

To get a picture of the proposed approach and the effect of τ on the approximation quality we conducted a simulation study. Consider over intervals $[-1,0]$ and $[0,1]$ polynomials $P_{\mathbf{a}}(x) = 1 - x + x^3$ and $P_{\mathbf{b}}(x) = 1 - x + 0.5x^3$, which at $x = 0$ fulfill the conditions $P_{\mathbf{a}}(0) = P_{\mathbf{b}}(0)$, $P'_{\mathbf{a}}(0) = P'_{\mathbf{b}}(0)$ and $P''_{\mathbf{a}}(0) = P''_{\mathbf{b}}(0)$. Based on $P_{\mathbf{b}}(x)$ and $x_{i,10} = i/10, i = \overline{1,10}$, random data $\widetilde{y}_{i,N}, i = \overline{1,10}$ were generated with $\varepsilon_{i,10} \sim \mathcal{N}(0, 0.05^2)$, see the first row of Table 1. For the sake of simplicity we did not

$\widetilde{y}_{i,10}$	0.91	0.76	0.77	0.65	0.50	0.51	0.46	0.44	0.50	0.42
b_i	7.80	-3.95	1.87	0.98	0.28	0.46	0.48	0.48	0.53	0.48

Tab. 1. Noisy ordinates and the estimate ($\tau = 0.01$) sequence b_i, $i = \overline{1, 10}$.

simulate ordinates for $P_{\mathbf{a}}(x)$. For estimating $b \equiv b_3$ using the recursion (17) we simply supposed that $P_{\hat{\mathbf{a}}}(x) \equiv P_{\mathbf{a}}(x)$ and so $\hat{I}_{i,10} \equiv I_{i,10}$. We remind that in our model (13)

$$\widetilde{y}(x) = I^{\mathcal{R}}(x) + Z(x)b + \varepsilon,$$

the incomplete interpolation $I^{\mathcal{R}}(x) = \sum_{i=0}^{2} p_i(x) P_{\hat{\mathbf{a}}}((i-2)\tau$ plays a threefold role: it ensures reparameterization, a quasi smooth transition between neighboring segments and a numerical saving: three parameters from four should not be estimated, but simply computed based on $P_{\hat{\mathbf{a}}}(x)$.

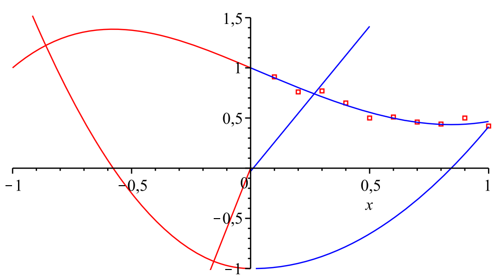

Fig. 1.

The recursive schema (17)

$$\begin{aligned} b_0 &= 0, \\ b_i &= b_{i-1} + \frac{Z_{i,10}}{\zeta_i}(\tilde{y}_{i,10} - \hat{I}_{i,10} - Z_{i,10} b_{i-1}), \ \ i = \overline{1,10}, \end{aligned}$$

where $Z_{i,10} = (x - 2\tau)(x - \tau)x, \quad \tau = 0.01, \quad \zeta_i = \sum_{k=1}^{i} Z_{k,10}^2$ and
$\hat{I}_{i,10} = 5099.96(\frac{i}{10} + 0.01)\frac{i}{10} - 10099.99(\frac{i}{10} + 0.02)\frac{i}{10} + 5000(\frac{i}{10} + 0.02)(\frac{i}{10} + 0.01)$,
produces the estimates $b_i, i = \overline{1, 10}$ for $b \equiv b_3$, see the second row of Table 1 rounded to two decimal digits. Hence the leading coefficient 0.5 of $P_{\mathbf{b}}(x)$ is estimated by $b_{10} = 0.48$. The approximant (16) with
$I^{\hat{\mathcal{R}}}(x) = 5099.96(x + 0.01)x - 10099.99(x + 0.02)x + 5000(x + 0.02)(x + 0.01)$
equals

$$\begin{aligned} P_{\hat{\mathbf{b}}}(x) &= I^{\hat{\mathcal{R}}}(x) + 0.4815(x + 0.02)(x + 0.01)x \\ (&= 1 - 1.0001037x - 0.01555x^2 + 0.4815x^3). \end{aligned}$$

τ	0.01	0.001	0.0001
b_{10}	0.4815	0.4658	0.4638
$P'_{\hat{\mathbf{b}}}(0)$	-1.0001	-1.0000	-1.0000
$P''_{\hat{\mathbf{b}}}(0)$	-0.0311	-0.004	0

Tab. 2. The impact of τ on the derivatives in the shared point zero.

τ	0.01	0.001	0.0001
Sample mean of $\lvert\hat{b}_{10} - 0.5\rvert$	0.02	0.002	0.0003
Sample variance of $\hat{b}_{10}$	0.0012	0.0012	0.0012

Tab. 3. Average deviations from $b_3 \equiv 0.5$.

Figure 1 depicts in addition to $P_{\mathbf{a}}(x)$, $\{\tilde{y}_{i,N}\}$ and $P_{\hat{\mathbf{b}}}(x)$ their first two derivatives, too. The plot of derivatives enable the visual check of smooth transition at the share point. According to the quasi smooth condition (11) $P_{\mathbf{a}}(0) = P_{\hat{\mathbf{b}}}(0)$, and $P'_{\mathbf{a}}(0) = P'_{\hat{\mathbf{b}}}(0) + o(\tau)$, $P''_{\mathbf{a}}(0) = P''_{\hat{\mathbf{b}}}(0) + o(\tau)$. Since $P_{\mathbf{a}}(0) = 1$, $P'_{\mathbf{a}}(0) = -1$ and $P''_{\mathbf{a}}(0) = 0$, the values $o(\tau)$ can be assessed from Table 2.

Table 2 was constructed with various τ values. As we can see, the smaller the τ is, the more accurate are the derivatives in zero computed from $P_{\hat{\mathbf{b}}}(x)$, that is the consequence of the LS method. The second derivative is more sensitive to the value of τ, however in accordance with the quasi smooth condition $P''_{\mathbf{b}}(0) = P''_{\mathbf{a}}(0) + o(\tau)$. Mention should be made that smaller τ does not imply more correct decimal places in b_i, see the second row of Table 2. Therefore we analyzed as well as the impact of τ on the average convergence of the estimate sequence. Table 3 contains three sample means $|0.5 - \hat{b}_{10}|$ of the difference $|0.5 - b_{10}|$ and the corresponding sample variances, each computed from 10000 random samples of length 10 with three different τ. We see that for $\tau = 10^{-k}$, $k = 2, 3, 4$, the recursive formula (17) returns in average $k - 1$ correct decimal places. The sample variances practically do not depend on the given values of τ. This result is in accordance with the theoretical value of the variance of b_{10} computed based on (21) or Theorem 4.2, $Db_{10} = \frac{\sigma^2}{\zeta_{10}} \approx 0.00118$, where $\sigma = 0.05$.

We can conclude based on the simulations that smaller τ implies smaller deviations of derivatives at the shared point $x = 0$, however smaller deviations of b_i from the estimated leading parameter only in average.

6. CONCLUSION

The paper introduced a new cubic approximation model based on a polynomial representation formula in which the key role is played by three reference points. Its main results are a two-part recursive smoothing scheme providing a smooth approximant for two neighboring local polynomials and the asymptotic normality of the second model's parameter from the two-part scheme. Our goal is to apply the scheme in knot detection algorithms to model the trend of signals.

Our approach to smoothing data can be generalized in several ways. Instead of the considered model one can also construct a two-part scheme with higher polynomial degree. The presented technique can be generalized as well as to smoothing data in space by polynomial surfaces.

The new approach is not limited to two polynomials and three reference points. Local and global piecewise smoothing schemas can be elaborated with any reasonable number of reference points and polynomial degree. However it is out of the scope of this work and is the task of the near future.

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 August 24, 2011)

REFERENCES

[1] N. D. Dikoussar: Adaptive projective filters for track finding. Comput. Phys. Commun. *79* (1994), 39–51.

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

[3] N. D. Dikoussar and Cs. Török: Automatic knot finding for piecewise-cubic approximation. Mat. Model. *T-17* (2006), 3.

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

[5] S. Haykin: Adaptive Filter Theory. Prentice Hall, 2002

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

[7] Ch. H. Reinsch: Smoothing by spline functions. Numer. Math. *10* (1967), 177–183.

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

[9] B. D. Ripley: Pattern Recognision and Neural Networks. Cambridge University Press, 1996.

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

[11] M. T. Wasan: Stochastic Approximation. Cambridge University Press, 2004.

Martina Révayová, Kybernetika s.r.o., Orgovánová 4, 040 11 Košice. Slovak Republic.
e-mail: martina.revayova@gmail.com

Csaba Török, Institute of Computer Science, UPJŠ, Jesenná 5, 040 01 Košice. Slovak Republic.
e-mail: csaba.torok@upjs.sk