International Scholarly Research Network
ISRN Applied Mathematics
Volume 2011, Article ID 321683, 10 pages
doi:10.5402/2011/321683

Research Article

Approximations of Time Series

M. Brackstone[1] and A. S. Deakin[2]

[1] *Department of Surgery, University of Western Ontario, London, Canada N6A 5B7*
[2] *Department of Applied Mathematics, University of Western Ontario, London, Canada N6A 5B7*

Correspondence should be addressed to A. S. Deakin, asdeakin@uwo.ca

Received 14 June 2011; Accepted 18 July 2011

Academic Editors: A. J. Kearsley, M. Tabata, and F. Tadeo

A method is proposed to approximate the main features or patterns including interventions that may occur in a time series. Collision data from the Ontario Ministry of Transportation illustrate the approach using monthly collision counts from police reports over a 10-year period from 1990 to 1999. The domain of the time series is partitioned into nonoverlapping subdomains. The major condition on the approximation requires that the series and the approximation have the same average value over each subdomain. To obtain a smooth approximation, based on the second difference of the series, a few iterations are necessary since an iteration over one subdomain is affected by the previous iteration over the adjacent subdomains.

1. Introduction

Graduated licensing system (GLS) is a method of gradual exposure of young novice drivers into the driving environment, allowing them to obtain initial experience with driving under supervision, followed by more independent driving under higher-risk circumstances [1]. This model was widely incorporated into driver licensing programs across the US and Canada as well as other countries over the 1990s. Most of these programs have incorporated similar restrictions into their initial phases [2]. These include driving with supervision, restricted driving at night, limited teenage passengers, and zero blood alcohol level while driving. This method has had limited long-term evaluation in North America, but long-term followup in New Zealand suggested a reduced but persistent long-term reduction in young driver collisions as a result of its implementation [3]. The collision data for Ontario drivers, around the time of the introduction of the GLS ([2], p. 126), illustrate the variety of approximations of time series that are possible.

There are many practical techniques for smoothing a time series [4]. The smoothed value at a point is a weighted average involving the elements in the series that are within a local window about the point. One way to generate the weights in a moving average involves a local polynomial approximation of order 3 (or 5) where the window includes 5 (or 7 et cetera) points. The weights are then defined by regression. Another approach defines the weights in terms of an appropriate kernel, and this method applies more generally to bivariate data [5]. The advantages of local estimates compared with global estimates are discussed in [6].

The first step in the computational process involves the partition of the domain of the time series into subdomains. The subdomains are then labelled as odd-numbered or even-numbered. Iterations are then performed over the odd numbered subdomains followed by the iterations over the even-numbered subdomains. This process is numerically efficient since the iterations over one set of subdomains update the boundary conditions for the iterations over the remaining subdomains [7]. To determine a smooth and accurate approximation, this set of iterations is repeated a few times.

This paper is organized as follows. Section 2 describes the form of the approximations, the partition of the time series, and the minimization over the subdomains of the partition. In Section 3, the equations for the approximation are derived, and some computational details are given in Section 4. Under certain assumptions, approximations of time series with variable spacing are possible. The time series and their approximations are presented in Section 5, and an example outlines the approach for step level changes, missing data, and outliers in Section 6. In Section 7, the approximation over a subdomain is determined by a fourth-order polynomial and a straight line. Finally, guidelines for the application of the proposed approach to a time series are outlined in "Concluding remarks."

2. The General Model

The equation for the general model for the approximation of the time series $\{Z_t \mid t = 1, \dots, N\}$ is

$$Z_t = \sum_{k=0}^{n-1} \left(O_t^k + Q_t^k \right) + R_t^n, \tag{2.1}$$

where $Q_t^0 = T_t + M_t$ and $R_t^n = A_t^n W_t^n$. In these equations, O_t^k is the term for outliers (if present); Q_t^k is the kth approximation; T_t is a trend and includes the level changes; M_t is a nonperiodic oscillatory function; R_t^n is the remainder; A_t^n is a measure of the variation of the remainder. A restriction on the approximations requires that the root-mean-square (RMS) value of the remainder is a decreasing sequence with increasing n. The form of (2.1) is similar to the asymptotic expansion of a function that contains a small positive parameter ([8], p. 1–4).

The partition of the domain of the time series $\{Z_t\}$, henceforth denoted by Z_t, is chosen in order to accurately approximate possible patterns in the time series. Let $P = \{E_k \mid k = 1, \dots, M\}$ be a partition of $[1, N]$ where the nonoverlapping subdomains are $E_k = \{t_k - n_k + 1, \dots, t_k\}$, $n_k \geq 1$, where n_k is the number of elements in the kth subdomain E_k. The overlapping subdomains, over which the iterations are computed, are defined as $E_k^o = \{t_k - n_k, E_k, t_k + 1\}$. For $k = 1$, $E_1^o = \{0, E_1, n_1 + 1\}$ and for $k = M$, $E_M^o = \{N - n_N, E_N, N + 1\}$ so that the overlapping subdomains are defined over the interval $[0, N + 1]$.

An approximation Q_t of a time series Z_t, where $Z_t = Q_t + R_t$, is determined by a few iterations I_t^n starting with $I_t^0 = Z_t$. Once the desired accuracy is obtained, the last iteration is defined as Q_t. All iterations and the approximation Q_t along with the remainder R_t satisfy the following properties.

(1) An iteration over E_k has the same average value a_k as the time series

$$\sum_{t \in E_k} Z_t = \sum_{t \in E_k} I_t^n = a_k, \tag{2.2}$$

and, hence, the average value of the remainder R_t is zero over this subdomain. If Z_t is a measure of "energy" in the process, then Q_t conserves energy over each subdomain in the partition. In the particular case $n_k = 1$, $E_k = \{t_k\}$ and $t = t_k$ is a fixed point for the approximation so that $Q_t := Z_t$ at $t = t_k$.

(2) The measure of smoothness of the iterations at time t for the nth iteration is defined by $\delta_t^n = I_{t+1}^n + I_{t-1}^n - 2I_t^n$ which is the second difference of I_t^n at time t. The norm on E_k^o is defined as the RMS value

$$\Delta_k^n = \left(\frac{\sum_{t \in E_k} \left(\delta_t^n\right)^2}{n_k} \right)^{1/2}. \tag{2.3}$$

Provided that the number of elements in E_k is greater than 1, then the condition that Δ_k^n has a minimum value is imposed. I_t^n is required for $t = t_k - n_k$ and $t = t_k + 1$ in E_k^o to determine δ_t^n at the endpoints of E_k. These two values of I_t^n are the boundary conditions for the minimization on E_k.

(3) For most of the examples presented in this paper, $t = 1$ and $t = N$ are fixed points so that $Q_1 = Z_1$ ($n_1 = 1$) and $Q_M = Z_N$ ($n_M = 1$). These values provide the boundary conditions for the minimization over the adjacent subdomain. For the general case where $n_1 > 1$ in E_1 ($n_M > 1$ in E_M), one of the boundary conditions is missing in E_1^o (E_M^o) so that an external boundary condition is required as described in the last paragraph of Section 3.

(4) In some cases there are two or more approximations over one or more subdomains and a criterion is required to choose the best approximation. From (2.1), $R_t^k = Q_t^k + R_t^{k+1}(R_t^0 := Z_t)$ where Q_t^k is determined from R_t^k over a partition P_k. For the example in Section 5, the simplest case occurs when P_k is a refinement of P_{k-1}; that is, $P_k = \cup P_k^\ell$ where P_k^ℓ covers E_ℓ in P_{k-1}. Then an approximation over E_ℓ is $Q_t^k = 0$ and the other is defined by P_k^ℓ. Let the RMS value of the remainder R_t^{k+1} over P_k^ℓ be denoted by S^{k+1}, and S^k is the RMS value over E_ℓ in R_t^k. The approximation defined by P_k^ℓ is a significant improvement if the ratio $S^{k+1}/S^k \leq \epsilon$ for a chosen value of ϵ. As shown in Section 7, an upper bound for ϵ takes on values between 0.75 and 0.9. For the example in Section 6 involving an outlier, there are two approximations for Q_t^0.

(5) The magnitude A_t^n of the remainder R_t^n is defined to be the RMS value of the remainder over each subdomain in a partition, and this definition implies that the RMS value of the series W_t^n in (2.1) is equal to 1. For the example presented in Section 5, the subdomains are uniform with 12 elements.

3. Mathematical Details

Given the iterates I_t^{n-1} and δ_t^{n-1} on E_k^o, the iterates I_t^n for $t \in E_k$ are computed such that the sum of squares of δ_t^n has a minimum value. For the moment, the first and last interval E_1

and E_M are excluded. The following variables are required to set up the equations for the minimization over the subdomain:

$$X_k^n = \left[\delta_{t_k-n_k+1}^n, \dots, \delta_{t_k}^n\right]', \qquad Y_k^n = \left[I_{t_k-n_k+1}^n, \dots, I_{t_k}^n\right]',$$
$$B_k^n = \left[I_{t_k-n_k}^n, 0, \dots, 0, I_{t_k+1}^n\right]' \tag{3.1}$$

are $n_k \times 1$ matrices and a prime on a matrix indicates the transposed matrix. From the definition of δ_t^n in Section 2, these matrices are related by $X_k^n - AY_k^n = B_k^n$, where A is a $n_k \times n_k$ tridiagonal symmetric matrix with elements

$$\{(-2,1); (1,-2,1); \dots; (1,-2,1); (1,-2)\}, \tag{3.2}$$

where -2 is on the main diagonal. The equations for the iterations are obtained by replacing B_k^n with B_k^{n-1}. Since the sum of I_t^n for $t \in E_k$ is a constant for all n, then

$$H'Y_k^n = \sum_{t \in E_k} Z_t = a_k n_k, \tag{3.3}$$

where $H = [1, 1, \dots, 1]'$, and a_k is the average value of Z_t over E_k. The condition on the sum in terms of X_k^n is $E'(X_k^n - B_k^{n-1}) = a_k n_k$, where E is the solution of $AE = H$. Thus, $E'X_k^n = n_k \rho_k^{n-1}$ where

$$\rho_k^{n-1} = a_k - \frac{\left(I_{t_{k-1}}^{n-1} + I_{t_k+1}^{n-1}\right)}{2}, \tag{3.4}$$

and $E[1] = E[n_k] = -n_k/2$ (Section 7). The solution for X_k^n such that $(X_k^n)'X_k^n$ has a minimum is $X_k^n = \rho_k^{n-1} G^2(n_k) E$, where $G^2(n_k) = n_k/(E'E)$. Finally, $\Delta_k^n = |\rho_k^{n-1}| G(n_k)$, and the solution Y_k^n is

$$Y_k^n = I_{t_k-n_k}^{n-1} V_1 + I_{t_k+1}^{n-1} V_2 + \rho_k^{n-1} V_3, \tag{3.5}$$

$$V_1[i] = 1 - \frac{i}{(n_k+1)}, \qquad V_2[i] = \frac{i}{(n_k+1)}, \qquad AV_3 = G^2(n_k). \tag{3.6}$$

The iteration $I_t^n = Y_k^n[i]$, where $i = 1, \dots, n_k$ and $t = t_k - n_k + 1, \dots, t_k$ in E_k, respectively. For a given n_k, E and then V_3 are uniquely determined, and V_3 is computed in advance for all of the subdomains that occur in a time series.

3.1. External Boundary Conditions

If $t = 1$ is not a fixed point, then the subdomain E_1^o requires a boundary condition. Here are three possible external boundary conditions to impose at $t = 0$ that can be used to reflect the possible behavior of the time series near the endpoint.

(1) $I_0^{n-1} = I_1^{n-1}$. The slope of the tangent is zero at $t = 1$.

(2) I_0^{n-1} is defined so that $\rho_1^{n-1} = 0$ in (3.4). This condition implies that the approximation over E_1 is a segment of a straight line.

(3) An iterative process is used to obtain Z_0 such that the RMS value of the remainder over the adjacent subdomain(s) has a minimum value.

Similarly, if $t = N$ is not a fixed point, then the external boundary conditions are obtained by replacing I_0^{n-1}, I_1^{n-1}, ρ_1^{n-1}, Z_0 with I_{N+1}^{n-1}, I_N^{n-1}, ρ_M^{n-1}, Z_{N+1}, respectively.

4. Computational Aspects

For a time series Z_t and a partition P, there is a related series defined by $\overline{Z}_t = a_k$ for $t \in E_k$, where a_k is the average value of Z_t for $t \in E_k$. This property holds for all of the approximations in this paper,

The approximations for a time series Z_t and the averaged time series $\overline{Z}_t$ are the same to the desired accuracy provided that the same partition and the same external boundary conditions (if any) are applied.

Consequently, any time series with variable spacing can be approximated provided that the estimates of the average values of the time series over the subdomains are adequate.

The approximation for the averaged series is employed especially for larger subdomains ($n_k \approx 12$ or more). The efficiency of the computations is increased if the boundary conditions in the first set of iterations (even and odd) are the average of the four values of the series that straddle the subdomains E_{k-1} and E_k. The averaged series was used in all computations, although the approximation obtained from Z_t may be more efficient in special cases.

It is convenient to introduce another notation to represent a partition: $P = \{n_1, n_2, \ldots; \ldots; \ldots, n_M\}$, where the number of elements in the subdomains in the first block is $\{n_1, n_2, \ldots\}$ and in the last block by $\{\ldots, n_M\}$. These blocks are a convenient way to separate the seasons or a set of months. Also, the approximation Q_t^k obtained by iterating the time series n times, using the partition P_k, is denoted by $\wp_k^n\{R_t^k\}$ $(R_t^0 := Z_t)$. The number of iterations n is determined from the difference

$$\{D_t^n\} = \wp_k^n\left\{R_t^k\right\} - \wp_k^{2n}\left\{R_t^k\right\} \tag{4.1}$$

by imposing the condition that $\max(|D_t^n|) < L$. $L = 1$ in Figures 1 and 2, $L = 0.04$ in Figures 3 and 4, and $L = 0.01$ in Figure 5. All calculations in this paper were performed using Maple software [9].

5. Applications

Two time series, provided by the Ontario Ministry of Transportation ([2], p. 126), illustrate the approximations. The graph of the time series for the monthly accidents for young novice drivers is given in Figure 1 where the main feature here is the intervention that occurs at 52 months owing to the introduction of the GLS on April 1, 1994. The corresponding graph for all drivers is shown in Figure 3 where the sharp drop in the graph from the maximum in December/January to April, except for the last 2 years, is a strong feature of the series.

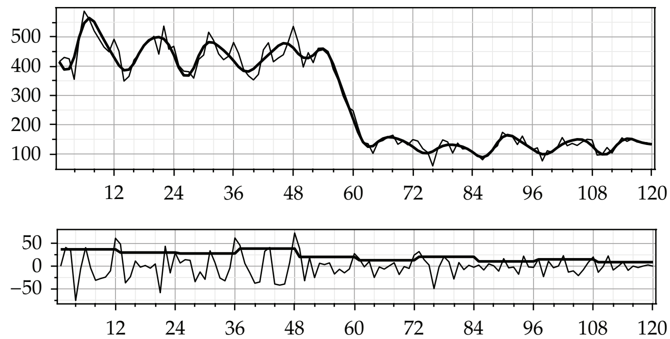

Figure 1: The upper graphs include Z_t (thin line), the accidents per month for young novice drivers, and the approximation Q_t^0 (thick line). The lower graphs represent the remainder $R_t^1 = Z_t - Q_t^0$ (thin line) and the amplitude of the remainder A_t^1 (thick line). The RMS value of the remainder is 25.

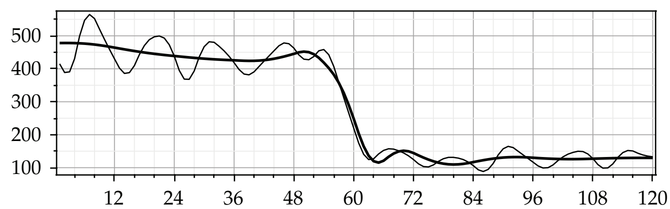

Figure 2: The graph of the trend T_t (thick line) and the approximation Q_t^0 (thin line) from Figure 1 for young novice drivers are shown.

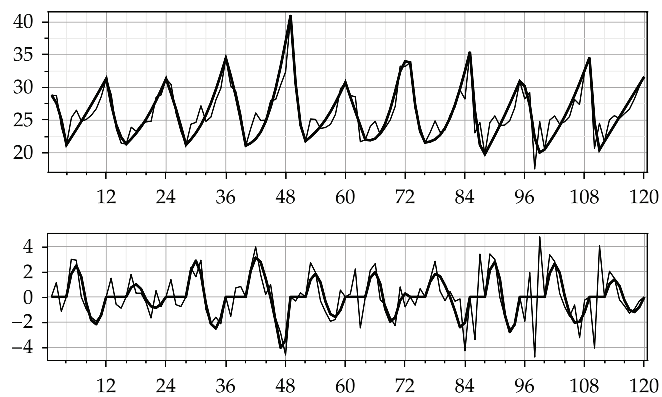

Figure 3: The upper graphs represent Z_t (thin line), the accidents (in thousands) per month for all drivers, and the approximation Q_t^0 (thick line). The RMS value of the remainder R_t^1 is 1.83. The lower graphs include the second approximation Q_t^1 (thick line) and the remainder R_t^1 (thin line).

In Figure 1, the uniform partition $P_0 = \{1,3;4;\ldots;4;3,1\}$, except for the first and last blocks, provides a smooth approximation $Q_t^0 = \wp_0^4\{Z_t\}$ and captures the intervention well. In this case, the first and last elements are fixed ($Q_1^0 = Z_1$ and $Q_{120}^0 = Z_{120}$). Since the sum of the elements in the remainder over E_k is zero, then $\wp_0\{R_t^1\} = \{0\}$. Other uniform partitions are possible where there are 3 or 6 elements in each subdomain. The approximation in the former case is slightly less smooth than Q_t^0 in Figure 1, and the RMS value of the remainder is 26. For

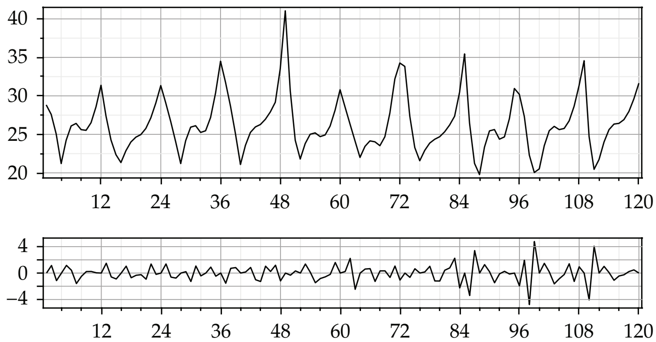

Figure 4: The upper graph is the approximation $Q_t^0 + Q_t^1$ (in thousands) for all drivers. The lower graph is the remainder R_t^2; the RMS value is 1.29.

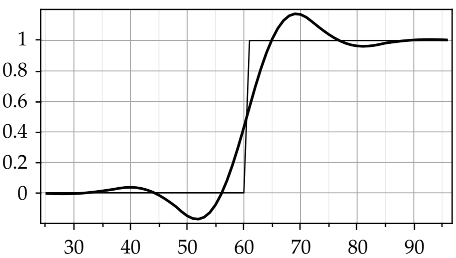

Figure 5: The trend T_t (thin line) of two segments of a series Z_t is $T_t = 0$ for $t \leq 60$ and $T_t = 1$ for $t \geq 61$; whereas, Q_t (thick line) is the trend of the series over [1,120].

the case of 6 elements, the RMS value of the remainder is 30. A more accurate approximation is obtained if the subdomains have two elements; however, this approximation has an angular appearance since it more closely approximates the time series.

In Figure 2, the approximation Q_t^0 of Figure 1 is expressed as a sum of a trend and an oscillatory series. The partition for the trend $T_t = \wp_T^6\{Q_t^0\}$ is $P_T = \{12, 12, 12, 12; 6, 6, 6, 6; 12, 12, 12, 12\}$. The external boundary condition implies that the tangent is horizontal at the endpoints of the series. The trend in this example is defined as a seasonal approximation of the time series where the subdomains contain 6 elements over the domain of the intervention. The remainder is the oscillatory series $M_t = Q_t^0 - T_t$.

In Figure 3, the points for January or December (plus one November) and April are fixed points for the approximation $Q_t^0 = \wp_0^3\{Z_t\}$. The partition is $P_0 = \{1, 2, 1, 7, 1; 3, 1, 7, 1; 3, 1, 7, 1; 3, 1, 8; 1, 2, 1, 7, 1; 3, 1, 6, 2; 1, 2, 1, 8; 1, 2, 1, 6, 1; 2, 2, 1, 8; 1, 2, 1, 7, 1\}$. The second approximation captures the increase in the number of accidents that occur in the summer months by approximating the remainder R_t^1 in $Z_t = Q_t^0 + R_t^1$ to obtain $R_t^1 = Q_t^1 + R_t^2$ where $Q_t^1 = \wp_1^3\{R_t^1\}$. The partition P_1 is a refinement of P_0 where 7 is replaced with 3, 4; 6 with 3, 3; 8 with 3,5. Consequently, the approximations Q_t^0 and $Q_t^0 + Q_t^1$ have the same average value over the subdomains of P_0.

The subdomains that are the same in the two partitions P_0 and P_1 are indicated by the intervals over which the approximation is zero in Figure 3. For the remaining intervals, the ratio of the RMS value of the remainder R_t^2 in Figure 4 to the RMS value of R_t^1 is equal to 0.49 so that this second approximation is significant. Furthermore, each of the ten segments of Q_t^1, excluding the segments in which the approximation is identically equal to zero, has a value between 0.40 and 0.52. The approximation $Q_t^0 + Q_t^1$ of Z_t and the remainder R_t^2 appear in Figure 4.

6. Level Changes, Missing Data, and Outliers

For a step level change between $t = \tau$ and $t = \tau + 1$, an approximation may not provide an adequate approximation for the time series in the subdomains on both sides of the step. For $t > \tau$, an external boundary condition is applied at $t = \tau$ such that the remainder of the approximation has a minimum RMS value. The same approach is applied to the series for $t < \tau + 1$. These ideas are illustrated in Figure 5 where the partition for the approximation $Q_t = \mathcal{P}^6\{T_t\}$ over [1,120] is $P = \{1, 11;\ 12; \ldots;\ 12;\ 11, 1\}$. The approximation exhibits a phenomenon that is similar to a Fourier series near a discontinuity in that the approximation overshoots on the right and undershoots to the left of the jump. The maximum and minimum of Q_t are 1.176 and −0.173. Moving away from the jump in either direction, the oscillations of $Q_t - T_t$ occur with rapidly decreasing amplitude. The details in item 4 of Section 2 are applied over the subdomains adjacent to $t = 60$ to choose between the two approximations of the trend. Interventions and level shifts, from an autoregressive moving average point of view, are presented in [10] and [11].

A simple example indicates the approach for a series that has a missing value or a possible outlier at $t = 6$. The series is $\{Z_t\} = \{0.6, -0.3, -0.5, -0.2, 0.4, Z_6, 1, 0.8, 0.9, 1.1, 1.0, 1.2\}$ where Z_6 is not defined in the case of a missing value. For both cases, the approximation Q_t^0 is determined for the partition $P = \{1; 4; 1; 5; 1\}$, where the value of the series is X at $t = 6$, such that the RMS value of the remainder is a minimum. A good initial estimate for X is the average value of the time series in a window about $t = 6$ [5]. Then an iterative process is started to obtain $X \approx 1.25$, as shown in Figure 6, and the RMS value of the remainder is 0.104. The smoothest approximation over P occurs for $X \approx 0.12$, where $\sum_2^{11} \delta_t^2$ has a minimum value. For the case of a possible outlier, $O_t^0 = 0$ for $t \neq 6$ in (2.1) and $O_6^0 = Z_6 - 1.25$ provided that the ratio of the RMS value of the approximation with $X = 1.25$ to the RMS value of the approximation under the assumption that Z_6 is not an outlier satisfies the condition in item 4 of Section 2.

7. Properties of Approximations

Approximations of Random Samples

The point of this exercise is to determine the ϵ in item 4 of Section 2 such that the only reasonable approximation for a series of random samples is Q_t^0 equal to the mean of the series. 12,000 random samples from the normal distribution with a mean of 0 and a standard deviation of 1 were generated using Maple to form 100 time series with 120 elements in each series. For each series, five approximations were determined where the subdomains of the uniform partition contained 3, 4, 6, 12, and 24 elements. The external boundary condition for the approximation is the condition of zero slope of the tangent. For each series, the ratio of the RMS value of the remainder for the approximation to the RMS value of the series were calculated, and the results are given in Table 1. The approximations corresponding to 24 and 12 elements are smooth and appear to reflect an underlying pattern in the series; whereas, for the cases 3 and 4, the approximations are contorted. An upper bound for ϵ is less than the minimum values in the range.

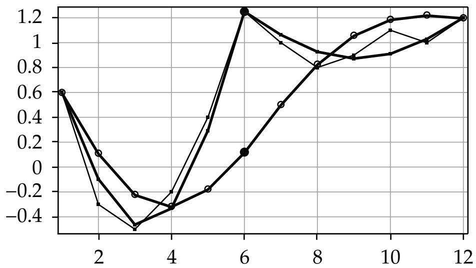

Figure 6: Graphs for the example are shown: $\{Z_t\}$ (thin line); approximation (thick line), $X = 1.25$ (large dot) at $t = 6$; smooth approximation, $X = 0.12$ at $t = 6$.

Table 1: The number of elements in the subdomains along with the mean, standard error (StE), and range of the ratios is shown.

Subdomain	Mean (StE)	Range
24	0.986 (0.013)	[0.94, 1.01]
12	0.967 (0.018)	[0.91, 1.00]
6	0.930 (0.023)	[0.87, 0.98]
4	0.888 (0.027)	[0.81, 0.96]
3	0.843 (0.035)	[0.76, 0.92]

Quartic Polynomial

The terms in the equation (3.5) for the approximation over E_k^o have a simple interpretation. For the first two terms in (3.5), $(i, V_1[i])$ and $(i, V_2[i])$ are points on straight lines. For the last term, Maple solves $AE = H$ for E and (3.6) for V_3 exactly; consequently, an accurate computation shows that $(i, E[i])$ and $(i, V_3[i])$ are points on the graph of a quadratic and a quartic polynomial, respectively.

To describe the properties of E and $V3$, it is necessary to change variables from t to s where $s = t - [t_k - (n_k - 1)/2]$ and $s = 0$ is the central point of the interval E_k^o. In the s variable, the boundary conditions are applied at the points $s = \pm s_o$, where $s_o = (n_k + 1)/2$. The equation of the quadratic polynomial $v(s, n_k)$ for E is defined by $d^2v/ds^2 = 1$ and $v = 0$ at $s = \pm s_o$ so that $v = (s^2 - s_o^2)/2$. For any integer i, $E[i]$ is equal to v at the corresponding value for s, and $E'E = ((2s_o)^5 - 2s_o)/120$. The equation for the quartic polynomial $u(s, n_k)$ for V_3, provided $n_k \geq 3$, is determined by $d^4u/ds^4 = G^2(n_k)$, where the roots of the equation for u are $s = \pm s_o$ and $s^2 = 5s_o^2 + 1$. Thus, $u = G^2(n_k)(s^2 - s_o^2)(s^2 - 5s_o^2 - 1)/24$. $G(n_k)$ is a measure of the smoothness of the approximation over E_k^o: G(2) = 1.0, G(3) = 0.59, G(4) = 0.39, G(6) = 0.21, G(12) = 0.062, and G(24) = 0.017.

Concluding Remarks

The major input for the approximation of a time series involves the partition of the domain. Initially a uniform partition is chosen and, if seasonal behavior is present in the series, a subset of the partitions cover the domain for the seasons. In general, as the length of the subintervals decreases, the approximation is less smooth and the accuracy of the approximation increases. The best approximation occurs at the point at which the approximation is acceptably smooth. The subintervals can be enlarged to determine a much smoother approximation that can be labelled as a trend while still respecting the seasonal aspects of the series; however, if an intervention is present, then some adjustment of the partition may be necessary in the region of the intervention. For time series with a well-defined local maximum or minimum, the approximation can be assigned the same value as the series by taking the partition to be a single point of the domain. For series with jumps and other complexities, examples are provided to suggest how to proceed in these cases.

An approach in the literature, as indicated in the introduction, defines the approximation at a point as a weighted average of the values of the values of the time series in a window about the point. This approach may smooth out interesting features in the time series and, if applied over a smaller intervals, the approximation will not be smooth. Since the proposed model is not based on regression, a comparison of the two approaches has not been considered.

References

[1] A. F. Williams and R. A. Shults, "Graduated driver licensing research, 2007-present: a review and commentary.," *Journal of Safety Research*, vol. 41, no. 2, pp. 77–84, 2010.

[2] M. Brackstone, *Proposal for impact evaluation of graduated licensing system on young drivers in Ontario*, M.S. thesis, Department of Epidemiology and Biostatistics, University of Western Ontario, Canada, 2008.

[3] J. D. Langley, A. C. Wagenaar, and D. J. Begg, "An evaluation of the New Zealand graduated driver licensing system," *Accident Analysis and Prevention*, vol. 28, no. 2, pp. 139–146, 1996.

[4] G. Janacek, *Practical Time Series*, Oxford University Press, New York, NY, USA, 2001.

[5] W. Härdle, *Applied Nonparametric Regression*, vol. 19 of *Econometric Society Monographs*, Cambridge University Press, New York, NY, USA, 1990.

[6] L. Keele, *Semiparametric Regression for the Social Sciences*, Wiley, Hoboken, NY, USA, 2008.

[7] B. F. Smith, P. E. Bjørstad, and W. D. Gropp, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equation*, Cambridge University Press, New York, NY, USA, 1996.

[8] J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics*, vol. 34 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1981.

[9] Maple software, version 13, Maplesoft, Waterloo, Canada.

[10] W. W. S. Wei, *Time Series Analysis: Univariate and Multivariate Method*, Addison Wesley/Pearson, Boston, Mass, USA, 2nd edition, 2006.

[11] R. S. Tsay, "Outliers, level shifts, and variance changes in time series," *Journal of Forecasting*, vol. 7, pp. 1–20, 1988.


Advances in Operations Research
Advances in Decision Sciences
Journal of Applied Mathematics
Algebra
Journal of Probability and Statistics
The Scientific World Journal
Hindawi
Submit your manuscripts at
http://www.hindawi.com
International Journal of Differential Equations
International Journal of Combinatorics
Advances in Mathematical Physics
Journal of Complex Analysis
Journal of Mathematics
Mathematical Problems in Engineering
Abstract and Applied Analysis
Discrete Dynamics in Nature and Society
International Journal of Mathematics and Mathematical Sciences
Journal of Discrete Mathematics
Journal of Function Spaces
International Journal of Stochastic Analysis
Journal of Optimization