Published in final edited form as:

Commun Stat Simul Comput. 2015 May ; 44(5): 1171–1186. doi:10.1080/03610918.2013.809103.

An Efficient Operator for the Change Point Estimation in Partial Spline Model

Sung Won Han^{1,*}, Hua Zhong², and Mary Putt³

¹Hoffmann-La Roche, Nutley, NJ, USA

²Department of Population Health, New York University, New York, NY, USA

³Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA

Abstract

In bio-informatics application, the estimation of the starting and ending points of drop-down in the longitudinal data is important. One possible approach to estimate such change times is to use the partial spline model with change points. In order to use estimate change time, the minimum operator in terms of a smoothing parameter has been widely used, but we showed that the minimum operator causes large MSE of change point estimates. In this paper, we proposed the summation operator in terms of a smoothing parameter, and our simulation study showed that the summation operator gives smaller MSE for estimated change points than the minimum one. We also applied the proposed approach to the experiment data, blood flow during photodynamic cancer therapy.

Keywords

Photodynamic therapy; change point; reproducing kernel Hilbert space; spline; nonparametric regression

1 Introduction

In many applications arising in biostatistics or bio-informatics, it can be of interest to estimate change points, starting and ending points of drop-down of the mean function, that give rise to longitudinal data. For example, the data from blood flow during photodynamic therapy (PDT) has a baseline curvature, and the change points related to the trend in the blood flow are related to the rate of tumor re-growth (Yu et al., 2005; Mesquita et al., 2011). Since the conventional parametric or nonparametric models are not appropriate for solving such a problem, Han et al. (2012) identified change points using a partial spline model based on the reproducing kernel Hilbert space-based (RKHS-based) spline.

In this paper, for the problem of the derivative change from a global trend under the unknown nonlinear baseline, a partial spline model with a reproducing kernel Hilbert

*Corresponding author. sungwonhan2@gmail.com.

space(RKHS)-based spline is considered. The partial spline can be treated as an unified framework including both linear and nonparametric model in terms of a baseline model. Laurent and Utreras (1986) discussed a partial spline model with a function which is smooth except for a discontinuity in a low-order derivative at a specific point. Shiao (1985) discussed various types of jump functions in several variables. Shiao et al. (1986) considered a jump function in two dimensions for modeling two dimensional atmospheric temperatures. Shiao (1985), Heckman (1986), and Shiao and Wahba (1988) studied the squared bias and variance of the coefficients of the function in the non-spline components of the function. For a comprehensive discussion of the partial spline, we refer to Chapter 6 in Wahba (1990).

The partial spline model with one derivative change point is mentioned by Wang (2011), who suggested a grid search algorithm and chose the estimated change point which minimizes GCV or AIC as a function of the smoothing parameter. However, the minimum operator in terms of the change time yields highly biased estimates of the change time even though the corresponding curve fitting is adequate. Thus, Han et al. (2012) proposed a new approach based on adaptive GCV, called aGCV, to allow more weight on the term for the generalized degrees of freedom, which led to much more accurate assessment of the change points. The method is computationally intensive and rapidly becomes unfeasible as the number of change points increases. Therefore, we propose a simpler and better approach to estimate change points, which is a way to find the time points minimizing $\sum_{\lambda} C(\lambda)$, where $C(\lambda)$ indicates criteria such as GCV or AIC. For the metric $C(\lambda)$, we consider several criteria such as *GML*, *AIC_c*, *CV* as well as *GCV* to show that the main contribution to estimate the change time depends on the summation operator, and not the specific criteria. The main difference of the proposed approach from the one suggested by Wang (2011) is that we use a summation operator rather than a minimum operator when we estimate the change time.

In Section 2, we discuss the partial spline model based on the RKHS-based spline and multiple change points, and explain how our model simultaneously detects change points in the global trend and captures the baseline curvature. In Section 3, we explain the smoothing parameter criteria based on the residual sum of square and degree of freedom for the estimated function. In Section 4, we explain our proposed method based on the summation operator to estimated change time. In Section 5, we compare the proposed method with other existing ones under several baselines and change times by using simulation studies. In Section 6, we apply them to several data sets as examples.

2 Statistical Methods

In this session, the partial spline model based on a RKHS-based spline with change points, which is mentioned in Wahba (1990) and Wang (2011), is described.

2.1 Partial spline model with change points

Let $Y = (y_1, \dots, y_n)^T$ be observations with corresponding times, $T = (t_1, \dots, t_n)^T$, in the time domain of $t \in [0, 1]$. For the i^{th} observation ($i = 1, \dots, n$), our model is:

$$y_i = f(t_i) + \sum_{k=1}^K \beta_k (t_i - \nu_k)_+ + e_i, \quad (1)$$

where $f(t_i)$ is an undetermined baseline function, ν_k is the k^{th} change point ($k = 1, \dots, K$), and β_k is the coefficient for change segment between ν_k and ν_{k+1} in the model, and e_i is the mean-zero error with $Var(e_i) = \sigma_e^2$.

The baseline function $f(t_i)$ in (1) is highly flexible. As described in Wang (2011) and Wahba (1990), for $f(t_i)$, we choose the cubic smoothing-spline model based on the reproducing kernel Hilbert space (RKHS), and it is written as:

$$f(t_i) = \theta_0 + \theta_1 t_i + \sum_{l=1}^n c_l \xi_l(t_i), \quad (2)$$

where $\xi_l(t) = P_1(t_i, t_l)$, and

$$P_1(t_i, t_l) = \int_0^1 (t_i - \gamma)_+ (t_l - \gamma)_+ d\gamma,$$

where $(t_i - \gamma)_+$ is the function indicating $(t_i - \gamma)I(t_i > \gamma)$. In Equation (2), the baseline has two parts: a linear term $\theta_0 + \theta_1 t$ and the remainder term $\sum_{l=1}^n c_l \xi_l(t_i)$. The form in Equation (2) allows flexibility in the baseline curvature.

When the change-points are known, plugging Equation (2) into Equation (1) yields the partial spline model for the mean function by

$$h(t_i) = \sum_{k=1}^K \beta_k (t_i - \nu_k)_+ + \theta_0 + \theta_1 t_i + \sum_{l=1}^n c_l \xi_l(t_i), \quad (3)$$

where $h(t_i) = E[y_i]$. When the change-points, ν_k , are known, parameters in $h(t_i)$ in Equation (3) can be estimated by minimizing the residual sum of squares subject to a penalty, i.e.,

$$\frac{1}{n} \sum_{i=1}^n \{y_i - h(t_i)\}^2 + \lambda \int_0^1 f^{(2)}(t) dt, \quad (4)$$

where $f^{(2)}(t)$ is the second order derivative with respect to time t , $\int_0^1 f^{(2)}(t) dt$ is a penalty function, and λ is the smoothing parameter.

2.2 Estimation by penalized least square

In this subsection, we explain the parameter estimation for the Equation (1) based on the penalized least square. Our parameter estimation and inference follows a partial spline model structure. Given fixed change times ν_k 's, the matrix form of Equation (1) is

$$Y = Z_\nu \beta + T\theta + \Sigma c + e, \quad (5)$$

where $Y = [y_1, y_2, \dots, y_n]^T$, Z_v is an $n \times K$ matrix with j^{th} row $\{(t_i - v_1)_+, (t_i - v_2)_+, \dots, (t_i - v_K)_+\}$, and $\beta = [\beta_1, \beta_2, \dots, \beta_K]^T$, T is an $n \times 2$ matrix with j^{th} row $\{1, t_i\}$, $\theta = [\theta_0, \theta_1]^T$, Σ is an $n \times n$ matrix with $\{P_1(t_i, t_l)\}_{j=1, l=1}^n$, $c = [c_1, c_2, \dots, c_n]^T$ and $e = [e_1, e_2, \dots, e_n]^T$.

Based on Equation (5) and the partial spline derivation in Wahba (1990) and Wang (2011), given λ and v_k , the penalized square for Equation (5) is

$$\frac{1}{n} \|Y - Z_v \beta - T \theta - \Sigma c\|^2 + \lambda c^T \Sigma c. \quad (6)$$

Let $V = [Z_v, T]$ and $d = [\beta^T, \theta^T]^T$. To estimate the parameters to minimize Equation (6), Wahba (1990) suggested QR decomposition (Dongarra et al., 1979). Let the QR decomposition of V be $V = [Q_1, Q_2][R^T, O^T]^T$. Q_1 , Q_2 , and R are $n \times (K+2)$, $n \times (n - K - 2)$, and $(K+2) \times (K+2)$. $Q = [Q_1, Q_2]$ is an orthogonal matrix, and R is upper triangular and invertible. O is a $(n - K - 2) \times (K+2)$ zero matrix. By Wahba (1990),

$$\hat{c} = Q_2(Q_2^T M Q_2)^{-1} Q_2^T Y, \quad (7)$$

$$\hat{d} = R^{-1} Q_1^T (Y - M c), \quad (8)$$

where $M = \Sigma + n\lambda I_{n \times n}$. The smoother matrix $H_{\lambda, v}$ such that $\hat{Y} = H_{\lambda, v} Y$ can be represented by

$$H_{\lambda, v} = I_{n \times n} - n\lambda Q_2(Q_2^T M Q_2)^{-1} Q_2^T. \quad (9)$$

The above estimate can be obtained given that λ and v are fixed. The grid search algorithm to select λ and v is proposed by Wang (2011), which uses double minimum operators in terms of change time, v , and the smoothing parameter, λ , by

$$\min_{\nu} \min_{\lambda} C(\nu, \lambda), \quad (10)$$

where $C(\nu, \lambda)$ is a model selection criterion such as GCV. Equation (10) is different from the conventional criterion, $\min_{\lambda} C(\lambda)$ since it has the additional operator $\min_{\nu}$ in terms of ν .

For the first step to solve Equation (10), we search the possible change time points on the observation times. Given each combination in ν , find λ^* which minimizes $C(\nu, \lambda)$. Then, find ν^* which minimizes $C(\nu, \lambda^*)$. Thus, the optimal ν^* is on the observation times. If necessary, we search neighbors of ν^* based on finer partitions of the previous grid to find the minimum of $C(\nu, \lambda)$. For example, if $K = 2$ and $\nu^* = (t_{j_1}, t_{j_2})$, set the partitions in the box by $[t_{j_1-1}, t_{j_1+1}] \times [t_{j_2-1}, t_{j_2+1}]$ and calculate $C(\nu, \lambda)$. Practically, it does not improve the criterion $C(\nu, \lambda)$ since there is no observation between the adjacent observation times of (t_{j_1}, t_{j_2}) , which may cause little improvement of $C(\nu, \lambda)$.

3 Smoothing Parameter Selection Criteria

In this section, we briefly review the criterion $C(\nu, \lambda)$ in (10) to select the smoothing parameter.

3.1 Criteria based on RSS and DF

For the selection of a smoothing parameter, the trade-off between goodness of fit and model complexity is often considered. The most common measure[s] for the goodness of fit is the residual sum of squares:

$$RSS(\lambda) = \sum_{i=1}^n (y_i - \hat{h}(t_i))^2 = \|(I - H_{\nu,\lambda})Y\|^2,$$

where $H_{\nu,\lambda}$ is the linear smoother in (9). For model complexity, Ye (1998) defined generalized degrees of freedom (DF), and DF is represented by the smoothing matrix $H_{\nu,\lambda}$ (Ruppert et al., 2003; Wang, 2011) such that

$$DF(\lambda) = \text{tr}(H_{\nu,\lambda}).$$

$\text{tr}(H_{\nu,\lambda})$ can also be interpreted as the effective number of parameters used in the smoothing fit (Hastie and Tibshirani, 1990; Hurvich et al., 1998).

By using $RSS(\lambda)$ and $DF(\lambda)$, several well-known criteria for selecting a smoothing parameter are as follows.

- $GCV = \frac{RSS(\lambda)/n}{\left(1 - \frac{DF(\lambda)}{n}\right)^2}$
- $CV = \sum_{i=1}^n \left(\frac{\{(I - H_{\lambda,\nu})Y\}_i}{I - [H_{\lambda,\nu}]_{ii}} \right)^2$ (Hutchinson and de Hoog, 1985)
- $AIC = n \log \frac{RSS(\lambda)}{n} + 2DF(\lambda) + n + 2$ (Hurvich et al., 1998)
- $AIC_c = n \log \frac{RSS(\lambda)}{n} + n \left(\frac{2\{DF(\lambda) + 1\}}{n - DF(\lambda) - 2} + 1 \right)$ (Hurvich and Tsai, 1989)
- $GML = \frac{RSS(\lambda)}{[\det^+(I - H_{\nu,\lambda})]^{1/(n-m)}}$ (Wecker and Ansley, 1983; Wahba, 1985)

In this paper, our main goal is to estimate the change point well, but fitting the curve well does not guarantee good estimation of the change point. In our previous simulation study, we found that the conventional criteria allow large degrees of freedom, essentially under-smoothing of the models. Here, wrongly estimated change times still minimize the value of the criteria. Thus, in the next subsection, we explain an heuristic algorithm proposed by Han et al. (2012) to improve the estimation of the change time.

3.2 The adjusted GCV

In order to estimate the change time, we can use the above criteria, $C(\nu, \lambda)$, with the operator in (10) since GCV and AIC_c are well known to be good criteria for selecting a smooth parameter for a good data fitting in the spline model. However, if the partial spline contains

the change points as parameters, our pre-simulation study shows that GCV or AIC_c does not work well. Alternatively, Han et al. (2012) propose a new criteria, called adaptive Generalized Cross Validation ($aGCV$), where

$$\log aGCV = \log \left(\frac{RSS(\lambda)}{n} \right) - w \log \left(1 - \frac{DF(\lambda)}{n} \right). \quad (11)$$

$w - 2$ is a weight parameter determined from the data. In practice, minimizing GCV often leads to a model with high degrees of freedom. $aGCV$ sets a weight variable for the generalized degree of freedom that may exceed 2, and thus lead to a more highly smoothed model than that chosen by GCV .

Han et al. (2012) proposed an algorithm to choose a weight that allows substantially improved estimation of the change points. The approach considers the minimum value of $aGCV$ as a function of increasing values of w . At the value of w and λ , where the change points are accurately estimate, there is a sharp decrease in degrees of freedom that minimize $aGCV$ as a function of w .

The $aGCV$ from Han et al. (2012) works well if the baseline is reasonably smooth without fluctuation, and the variance of the curve is small. However, in real data, the baseline does not have such a smooth form and the noise size can also be large. In addition, this $aGCV$ is computationally intensive. For example, searching the grid of the weight w by 100 units for $aGCV$ takes 100 times longer than GCV . Therefore, we propose a new simple approach to estimate the change time in the next section.

4 Summation operators For Estimating Change Times

As we mentioned in the previous section, the $\min_{\nu} \min_{\lambda} C(\nu, \lambda)$ criteria often yields highly biased estimates of the change times. The reason is that the change times may contribute little to the overall value of $C(\nu, \lambda)$. As an example, consider Figure 1. The example is from one simulation with small curvature, the true change time $\nu=0.2$, and the true change size $\beta=-50$. The black curve indicates the $C(\nu, \lambda)$ values through different values of $\log(\lambda)$ given the true change time $\nu=0.2$. The minimum value in the $C(\nu, \lambda)$ curve is 0.1183 at $\log(\lambda)=-6.1$. However, given an incorrect change time $\nu = 0.92$, the minimum value in the $C(\nu, \lambda)$ curve (gray line) is 0.1097 at $\log(\lambda) = -5.9$.

This study suggest that the minimum operator rather than the criterion may be the source of the problem. Rather than the measure by

$$\min_{\nu} \min_{\lambda} C(\nu, \lambda), \quad (12)$$

we suggest the summation operator in terms of the smoothing parameter λ such that

$$\min_{\nu} \sum_{\lambda} C(\nu, \lambda) \quad (13)$$

in order to estimate ν . Then, given the estimate $\hat{\nu}$ from (13), we select the smoothing parameter λ by

$$\min_{\lambda} C(\hat{\nu}, \lambda). \quad (14)$$

The performance of our procedure is evaluated in terms of bias, variance, and mean square error (MSE). We found that the variance of the estimated change time can be significantly reduced by a summation operator, and suggested that the sensitivity of the change point estimation to noise in the baseline function results largely from the minimum operator, not the criteria themselves.

5 Simulation Study

In this section, we investigate the performance of the operator and criteria for estimating the change points in Equation (3). Motivated by the blood flow data in Han et al. (2012), to construct the baseline function, i.e. $f(t)$, we use the logistic model (Pinheiro and Bates, 2000) :

$$f(t_i) = \frac{\phi_1}{1 + \exp[-(t_i - \phi_2)/\phi_3]}, \quad (15)$$

where ϕ_1 indicates an asymptotic upper limit of simulated data, ϕ_2 the location parameter, and ϕ_3 the scale parameter. Varying ϕ_2 and ϕ_3 generates data with variable curvature. As in Han et al. (2012), we study the simulation under the two cases of the baseline $h(t_i)$: Case A with $\phi_1=100$, $\phi_2=-0.25$, and $\phi_3=0.15$, and Case B with $\phi_1=100$, $\phi_2=0$, and $\phi_3=0.1$. For the one change point cases, we add a change point at $\nu=0.2$ or 0.6 . We set the derivative change size β by $\beta=-50$ or -100 . Plots of the mean function with one change point appear in the upper panels of Figure 2, and for the two change point cases, where we add change points at $\nu_1=0.2$ and $\nu_2=0.6$, in the bottom panel of Figure 2. The simulation setting for the baseline and change patterns indicates that initially the baseline has small or large curvature, and at a later time, the baseline has no curvature. The change point is located in either the region of higher or lower curvature, or both.

5.1 Comparison of performance between summation and minimum operator

We compare the performance of two operators, minimum and summation, given widely used criteria such as GCV, GML, or AIC_c . We first investigate the efficiency of change point estimation. Table 1 shows $bias^2$, variance, and MSE for the estimated change time for the one change point case. As shown in the table, the MSE from the summation operator is significantly smaller than the MSE from the minimum operator among almost all criteria except for under the large curvature with $\nu = 0.6$. For example, in Table 1, under the small curvature with $\nu=0.2$, if the change size is -50 , the MSE from the minimum operator with GCV or CV is 0.0746 or 0.0790, respectively. On the other hand, the MSE's from the summation operator with those are 0.0003. Similarly, the MSE's of GML, AIC, or AIC_c from the minimum operator is over 0.01, but those from the summation operator are less than 0.0003.

The small MSE of the summation operator reflects the small variance in the estimation of the change time. In Table 1, the variance from the summation operator is consistently less than

0.001 except under small curvature with $v=0.6$ and change size $=-50$. However, most variances from the minimum operator are greater than 0.001. Unlike the variance difference, many biases from the minimum operator are similar to or even larger than those from the summation operator. For example, in Table 1, under small curvature with $v = 0.6$, if the change size is -100 , the $bias^2$ from the minimum operator with GCV, CV, or GML is <0.0001 , 0.0005 , or 0.0001 , respectively, but those from the summation operator are 0.0004 . Furthermore, under large curvature with $v = 0.6$ and a change size of -100 , the bias from the minimum operator tends to be smaller than that from the summation operator. Such a variance comes from the realization of noise, so we can see that the minimum operator is sensitive to noise. One interesting observation in Table 1 is that GML criteria show similarly good MSE no matter which operator is chosen. GML usually gives small degree of freedom, so it may not be much sensitive to noise.

The superiority of the summation operator over the minimum operator is clearly shown in the two change point cases, as shown in Table 2. For the two change point case, the variance and MSE from the summation operator is generally much smaller than that of the minimum operator. This trend tends to hold as well for the bias. For example, in Table 1, for the one change point case, under large curvature with $v = 0.6$, the $bias^2$ from the minimum operator tends to be smaller than those from the summation operator. In contrast, for the two change point case, the summation operator gives smaller $bias^2$ than the minimum operator (Table 2). Thus, The summation operator appears superior for estimating the change point when there are more change points, and especially when the baseline has a high curvature and the change point is on the region of the baseline where there is low curvature.

In addition, we investigated the MSE of the estimated curve $\hat{h}(t)$ between two operators. In contrast to the estimation of the change point, we found that the MSE's of estimated curve $\hat{h}(t)$ were similar between two operators. The difference is minor in comparison with that in terms of change time estimation.

5.2 Comparison between summation operator and aGCV

Han et al. (2012) proposed an algorithm to estimate the change points and the function at the change point based on a data-dependent form of GCV. The aGCV approach is also based, in part, on a minimum operator. Here, we compare the properties of aGCV with those of the summation operator.

For the one change point case, the MSE values from the summation operator with GCV are similar to or smaller than those from aGCV. When the change size is small, the summation operator tends to perform much better than aGCV in terms of MSE. For example, in Table 1, under large curvature with $v = 0.2$, if the change size is -50 , the MSE from the summation operator with GCV is 0.0006 , but the MSE from aGCV is 0.0253 . On the other hand, under the large curvature with $v = 0.6$ and change size $=-100$ from the one change point case, the MSE from the summation operator is 0.0236 while that from aGCV is 0.0396 , which are relatively close. It should be noted that for the one change point case, there are situations where aGCV yields smaller bias, in particular when the baseline curvature is large and the change point is a $v = 0.6$, or when the baseline curvature is small and $v = 0.2$.

For the two change point cases, all MSE as well as the bias term from the summation operator are similar to or smaller than those from aGCV. Thus, the summation operator appears superior for the two change point case.

5.3 Robustness of summation operator for larger noise

If the noise $\sigma_e^2=1$, the summation operator gives better estimation of change times than the minimum operator. Table 3 shows that for a large variance, $\sigma_e^2=2$, and one change point. The summation operator also outperforms the minimum operator. Similar results were observed for the two change point model (Results not shown). Generally, as the variance of noise becomes larger, the MSE of estimated change time also becomes larger since the variance of estimated change time increases.

Under $v = 0.2$ and change size $= -50$, we observed that the bias under $\sigma_e^2=1$ is larger than the bias under $\sigma_e^2=2$ (Results not shown). Thus, for most criteria, the MSE of estimated change time decreases as σ_e^2 increases. In most cases, the MSE from the minimum operator increases substantially, whereas the MSE from the summation operator shows little change. Even for GML criteria, which give good performance with minimum operator under $\sigma_e^2=1$, as noise σ_e^2 increases, MSE from minimum operator increases significantly, but the MSE from summation operator does not change much.

6 Example

The motivating example comes from observations of blood flow during photodynamic therapy (PDT), a treatment for cancer in solid tumors (Han et al. (2012)). Yu et al. (2005) found that the pattern of blood flow decline during the photodynamic therapy predicts the re-growth rate of the tumor. Thus, the detection of the starting and ending points of downward trend in blood flow during cancer treatment therapy is important. As examples, we apply three methods; minimum operator with GCV (GCV,min), summation operator (GCV,sum), and aGCV to two mice blood flow data.

The resulting plots appear in Figure 3. The upper part of Figure 3 shows the results from the data for Mouse A. The estimated change times from (GCV,min) are $\hat{v}_1=10.2$ minutes and $\hat{v}_2=13.6$ minutes. On the other hands, the estimate change time from (GCV,sum) and aGCV are same and $\hat{v}_1=18.7$ minutes and $\hat{v}_2=24.4$ minutes. The estimate from (GCV,min) looks poor, but other two seem reasonable. The $\log(\lambda)$ from (GCV,min), (GCV,sum), and aGCV are -10 , -8.1 , and -3.5 , respectively.

For Mouse B, the result from (GCV,sum) is more reasonable than the results from the other two operators. The estimated change times from (GCV,min) are $\hat{v}_1=0.6$ minutes and $\hat{v}_2=13.3$ minutes and those from aGCV are $\hat{v}_1=0.6$ minutes and $\hat{v}_2=12.2$ minutes. On the other hand, the estimated change times from (GCV,sum) are $\hat{v}_1=9.1$ minutes and $\hat{v}_2=12.5$ minutes. The $\log(\lambda)$ from (GCV,min), (GCV,sum), and aGCV are -8.9 , -7.3 , and -6.8 , respectively. As Figure 3 shows, (GCV,sum) shows a more reasonable change time estimate than (GCV,min) in both mice, and a more reasonable estimate than aGCV in Mouse B. The (GCV,min) and

aGCV depend on a specific choice of λ , whereas (GCV,sum) considers all choices of λ , the property of which appears to add robustness to the estimated change times.

7 Conclusion

In this paper, we describe a method for estimating change points in a partial spline model. We investigate the performance of the summation operator in terms of bias, variance, and MSE of the estimated change time by comparing it to those of the minimum operator or adjusted GCV (aGCV). The simulation study shows that the often recommended minimum operator has larger MSE of change point estimates than the summation operator. This summation operator approach gives similar performance as that from the computationally intensive aGCV. We also applied the proposed summation operator to the experimental data, blood flow during photodynamic cancer therapy, where it performs well. Since aGCV has high computational burden, but the summation operator approach is much simpler and often gives better performance, the summation operator is preferred.

Acknowledgements

We thank professor Theresa Busch at University of Pennsylvania, Philadelphia, PA, USA and Professor Rickson C. Mesquita at University of Campinas, Campinas, SP, Brazil for supporting the motivating data. The research is supported by NIH-NCI 5-P01-CA-087971.

References

- Dongarra, J.; Bunch, J.; Moler, C.; Stewart, G. Linpack Users' Guide. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1979.
- Han SW, Busch T, Mesquita RC, Putt M. A semi-parametric model for detecting change-points in blood flow. submitted. 2012
- Hastie, T.; Tibshirani, R. Generalized Additive Models. London: Chapman and Hall; 1990.
- Heckman N. Spline smoothing in a partly linear model. Journal of the Royal Statistical Society. Series B. 1986; 48:244–248.
- Hurvich CM, Tsai C-L. Regression and time series model selection in small samples. Biometrika. 1989; 76:297–307.
- Hutchinson MF, de Hoog FR. Smoothing noisy data with spline functions. Numerische Mathematik. 1985; 47:99–106.
- Hurvich CM, Simonoff JS, Tsai C-L. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. Journal of the Royal Statistical Society. Series B. 1998; 60:271–293.
- Laurent P, Utreras F. Optimal smoothing of noisy broken data with spline functions. Journal of Approximation Theory and its Applications. 1986; 2:71–94.
- Mesquita R, Putt M, Pole A, Han SW, Busch TM. Mouse strain affects the dynamics of tumor vascular response during PDT. PLoS One. 2011; 7:e37322. [PubMed: 22624014]
- Pinheiro, JC.; Bates, DM. Mixed-Effects Models in S and S-PLUS. New York, NY: Springer; 2000.
- Ruppert, D.; Wand, MP.; Carroll, RJ. Semiparametric regression. Cambridge, UK: Cambridge University Press; 2003.
- Shiau, J. Ph.D. thesis. Madison: Department of Statistics, University of Wisconsin; 1985. Smoothing spline estimation of functions with discontinuities.
- Shiau J, Wahba G, Johnson D. Partial spline models for the inclusion of tropopause and frontal boundary information. Journal of Atmospheric and Ocean Technology. 1986; 3:714–725.
- Shiau J, Wahba G. Rates of convergence of some estimators for a semiparametric model. Communications in Statistics - Simulation and Computation. 1988; 17:1117–1133.

- Wahba G. A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. *The Annals of Statistics*. 1985; 13:1378–1402.
- Wahba, G. *Spline models for observational data*. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1990.
- Wang, Y. *Smoothing Splines: Methods and Applications*. New York, NY: CRC Press, Taylor & Francis Group; 2011.
- Wecker WE, Ansley CF. The signal extraction approach to nonlinear regression and spline smoothing. *Journal of the American Statistical Association*. 1983; 78:81–89.
- Ye J. On measuring and correcting the effects of data mining and model selection. *Journal of the American Statistical Association*. 1998; 93:120–131.
- Yu G, Durduran T, Zhou C, Wang HW, Putt ME, Saunders HM, Sehgal CM, Glatstein E, Yodh AG, Busch TM. Noninvasive monitoring of murine tumor blood flow during and after photodynamic therapy provides early assessment of therapeutic efficacy. *Clinical Cancer Research*. 2005; 11:3543–3552. [PubMed: 15867258]

Figure 1.

GCV plots through $\log(\lambda)$ given the estimated change time: The black line indicates the GCV from the true change time, and the gray line indicates the GCV from the estimated change time. The vertical line indicates the optimal $\log(\lambda)$ giving the minimum GCV in each curve.

Figure 2.

The plots for the baseline with change points used in simulation study: The function with a single change-point is in the upper two panels, and the function with two change-points is in the bottom panel. The black line indicates the baseline with a change point with $\beta = -50$, and the gray line does with a change point with $\beta = -100$.

Figure 3.

Example of change point estimation in blood flow data using GCV criteria: The upper panel shows the blood flow data in Mouse A, and the bottom panel does in Mouse B. The left plots show the estimation from minimum operator, the middle plots do from summation operator, and the right plots do from aGCV. The two vertical lines indicate the estimated change times, and the curves through the data points indicate the estimated curve.

$bias^2$, variance, and MSE of the estimated change time, $\hat{v}$, depending on operator (sum or min) and criteria (i.e. GCV): one change point and noise $\sigma_e^2=1$.

Table 1

Small curvature		$v=0.2$						$v=0.6$					
Change size	measure	GCV	CV	GML	AIC	AIC_c	aGCV	GCV	CV	GML	AIC	AIC_c	aGCV
-50	min, $bias^2$	0.0111	0.0080	0.0003	0.1092	0.0053	0.0000	0.0036	0.0083	0.0012	0.0058	0.0050	0.0031
	min, variance	0.0635	0.0710	0.0110	0.0769	0.0477	0.0017	0.0484	0.0692	0.0095	0.0791	0.0503	0.0255
	min, MSE	0.0746	0.0790	0.0113	0.1861	0.0530	0.0017	0.0520	0.0775	0.0107	0.0849	0.0553	0.0286
	sum, $bias^2$	0.0002	0.0002	0.0000	0.0002	0.0002		0.0026	0.0026	0.0028	0.0022	0.0022	
	sum, variance	0.0002	0.0001	0.0000	0.0001	0.0001		0.0003	0.0003	0.0016	0.0004	0.0004	
	sum, MSE	0.0003	0.0003	0.0000	0.0003	0.0003		0.0030	0.0029	0.0045	0.0026	0.0026	
-100	min, $bias^2$	0.0044	0.0039	0.0000	0.1102	0.0010	0.0000	0.0000	0.0005	0.0001	0.0060	0.0002	0.0000
	min, variance	0.0395	0.0411	0.0002	0.0762	0.0228	0.0001	0.0162	0.0226	0.0003	0.0787	0.0086	0.0001
	min, MSE	0.0439	0.0450	0.0003	0.1863	0.0238	0.0001	0.0162	0.0230	0.0004	0.0847	0.0088	0.0001
	sum, $bias^2$	0.0000	0.0000	0.0000	0.0000	0.0000		0.0004	0.0004	0.0004	0.0003	0.0003	
	sum, variance	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000	
	sum, MSE	0.0001	0.0001	0.0000	0.0001	0.0001		0.0005	0.0005	0.0004	0.0003	0.0003	
Large curvature		$v=0.2$						$v=0.6$					
Change size	measure	GCV	CV	GML	AIC	AIC_c	aGCV	GCV	CV	GML	AIC	AIC_c	aGCV
-50	min, $bias^2$	0.0419	0.0367	0.0000	0.1084	0.0477	0.0006	0.0289	0.0389	0.1287	0.0050	0.0331	0.1185
	min, variance	0.0934	0.0928	0.0023	0.0800	0.0945	0.0246	0.0791	0.0828	0.0319	0.0800	0.0786	0.0413
	min, MSE	0.1353	0.1295	0.0023	0.1884	0.1422	0.0253	0.1079	0.1216	0.1607	0.0850	0.1117	0.1598
	sum, $bias^2$	0.0006	0.0005	0.0003	0.0004	0.0004		0.0828	0.0846	0.0838	0.0973	0.0973	
	sum, variance	0.0001	0.0001	0.0002	0.0001	0.0001		0.0001	0.0001	0.0005	0.0006	0.0006	
	sum, MSE	0.0006	0.0006	0.0005	0.0005	0.0005		0.0830	0.0848	0.0843	0.0980	0.0980	

-100	min, $bias^2$	0.0204	0.0210	0.0000	0.1063	0.0184	0.0005	0.0047	0.0049	0.0037	0.0059	0.0052	0.0070
	min, variance	0.0668	0.0857	0.0001	0.0811	0.0639	0.0154	0.0405	0.0486	0.0206	0.0822	0.0405	0.0326
	min, MSE	0.0872	0.1068	0.0001	0.1875	0.0823	0.0159	0.0451	0.0534	0.0242	0.0881	0.0457	0.0396
	sum, $bias^2$	0.0003	0.0002	0.0001	0.0001	0.0001		0.0234	0.0246	0.0257	0.0233	0.0233	
	sum, variance	0.0000	0.0000	0.0001	0.0000	0.0000		0.0002	0.0003	0.0000	0.0009	0.0009	
	sum, MSE	0.0003	0.0003	0.0002	0.0002	0.0002		0.0236	0.0249	0.0257	0.0241	0.0241	

Table 2

$bias^2$, variance, and MSE for each estimated change time, $\hat{v}_1$ and $\hat{v}_2$, depending on operator (sum or min) and criteria (i.e. GCV) (i.e. GCV): two change points and noise $\sigma_e^2=1$

Small curvature		v=0.2						v=0.6					
Change size	measure	GCV			GML			GCV			GML		
		GCV	CV	AIC	AIC _c	aGCV		GCV	CV	AIC	AIC _c	aGCV	
-50	min, $bias^2$	0.0045	0.0087	0.0001	0.0145	0.0040	0.0002	0.0045	0.0017	0.0002	0.0086	0.0014	0.0001
	min, variance	0.0385	0.0520	0.0001	0.0412	0.0283	0.0067	0.0341	0.0403	0.0075	0.0419	0.0234	0.0040
	min, MSE	0.0430	0.0607	0.0002	0.0557	0.0323	0.0069	0.0386	0.0420	0.0077	0.0504	0.0248	0.0040
	sum, $bias^2$	0.0000	0.0000	0.0000	0.0000	0.0000		0.0001	0.0001	0.0001	0.0001	0.0001	
	sum, variance	0.0001	0.0001	0.0001	0.0001	0.0001		0.0002	0.0002	0.0007	0.0002	0.0002	
	sum, MSE	0.0001	0.0001	0.0001	0.0001	0.0001		0.0003	0.0003	0.0008	0.0003	0.0003	
-100	min, $bias^2$	0.0012	0.0005	0.0001	0.0134	0.0001	0.0000	0.0003	0.0005	0.0000	0.0074	0.0001	0.0000
	min, variance	0.0189	0.0112	0.0001	0.0404	0.0070	0.0000	0.0202	0.0215	0.0001	0.0461	0.0110	0.0001
	min, MSE	0.0202	0.0117	0.0001	0.0538	0.0071	0.0001	0.0205	0.0220	0.0001	0.0535	0.0111	0.0001
	sum, $bias^2$	0.0000	0.0000	0.0001	0.0000	0.0000		0.0000	0.0000	0.0001	0.0000	0.0000	
	sum, variance	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0001	0.0003	0.0000	0.0000	
	sum, MSE	0.0000	0.0000	0.0001	0.0000	0.0000		0.0001	0.0001	0.0004	0.0001	0.0001	
Large curvature		v=0.2						v=0.6					
Change size	measure	GCV			GML			GCV			GML		
		GCV	CV	AIC	AIC _c	aGCV		GCV	CV	AIC	AIC _c	aGCV	
-50	min, $bias^2$	0.0189	0.0171	0.0003	0.0157	0.0099	0.0013	0.0162	0.0165	0.0086	0.0336	0.0839	
	min, variance	0.0785	0.0772	0.0005	0.0415	0.0647	0.0124	0.0686	0.0763	0.0248	0.0418	0.0601	0.0230
	min, MSE	0.0974	0.0942	0.0009	0.0572	0.0746	0.0137	0.0848	0.0928	0.0411	0.0505	0.0937	0.1069
	sum, $bias^2$	0.0001	0.0000	0.0000	0.0001	0.0001		0.0014	0.0015	0.0012	0.0012	0.0012	
	sum, variance	0.0000	0.0000	0.0001	0.0001	0.0001		0.0004	0.0004	0.0025	0.0005	0.0005	
	sum, MSE	0.0001	0.0001	0.0001	0.0001	0.0001		0.0018	0.0018	0.0037	0.0016	0.0016	

-100	min, $bias^2$	0.0144	0.0113	0.0000	0.0145	0.0100	0.0000	0.0010	0.0029	0.0000	0.0074	0.0022	0.0092
	min, variance	0.0506	0.0544	0.0001	0.0408	0.0511	0.0102	0.0462	0.0531	0.0026	0.0461	0.0472	0.0287
	min, MSE	0.0650	0.0657	0.0001	0.0553	0.0611	0.0103	0.0472	0.0560	0.0027	0.0535	0.0494	0.0379
	sum, $bias^2$	0.0000	0.0000	0.0000	0.0000	0.0000		0.0003	0.0003	0.0006	0.0002	0.0002	
	sum, variance	0.0000	0.0000	0.0001	0.0000	0.0000		0.0001	0.0001	0.0000	0.0001	0.0001	
	sum, MSE	0.0001	0.0000	0.0001	0.0000	0.0000		0.0003	0.0003	0.0006	0.0003	0.0003	

Table 3

MSE of the estimated change time, $\hat{v}$, depending on operator (sum or min) and criteria (i.e. GCV): one change point and noise $\sigma_e^2=2$

Small curvature								
v	Change size	measure	GCV	CV	GML	AIC	AIC_c aGCV	
0.2	-50	min, MSE	0.0629	0.0516	0.0509	0.1681	0.0389	0.0007
		sum, MSE	0.0007	0.0006	0.0005	0.0007	0.0007	
	-100	min, MSE	0.0490	0.0320	0.0306	0.1681	0.0249	0.0002
		sum, MSE	0.0002	0.0001	0.0000	0.0001	0.0001	
0.6	-50	min, MSE	0.0961	0.1088	0.0121	0.0904	0.0972	0.0482
		sum, MSE	0.0037	0.0040	0.0108	0.0037	0.0037	
	-100	min, MSE	0.0496	0.0467	0.0027	0.0904	0.0478	0.0049
		sum, MSE	0.0006	0.0006	0.0005	0.0005	0.0005	
Large curvature								
v	Change size	measure	GCV	CV	GML	AIC	AIC_c aGCV	
0.2	-50	min, MSE	0.1452	0.1136	0.0163	0.1253	0.1160	0.0025
		sum, MSE	0.0008	0.0008	0.0005	0.0008	0.0008	
	-100	min, MSE	0.1319	0.1151	0.0044	0.1253	0.1119	0.0075
		sum, MSE	0.0004	0.0004	0.0000	0.0004	0.0004	
0.6	-50	min, MSE	0.1430	0.1501	0.1568	0.0938	0.1480	0.1883
		sum, MSE	0.0808	0.0827	0.0653	0.0877	0.0877	
	-100	min, MSE	0.0859	0.1037	0.0582	0.0929	0.0865	0.1460
		sum, MSE	0.0236	0.0244	0.0272	0.0234	0.0234	