



An algorithm for a constrained P-spline

Rosanna Campagna¹ · Serena Crisci¹ · Gabriele Santin² · Gerardo Toraldo¹ · Marco Viola³

Received: 18 February 2024 / Accepted: 2 May 2025 / Published online: 21 May 2025
© The Author(s) 2025

Abstract

Regression splines are largely used to investigate and predict data behavior, attracting the interest of mathematicians for their beautiful numerical properties, and of statisticians for their versatility with respect to the applications. Several penalized spline regression models are available in the literature, and the most commonly used ones in real-world applications are P-splines, which enjoy the advantages of penalized models while being easy to generalize across different functional spaces and higher degree order, because of their discrete penalty term. To face the different requirements imposed by the nature of the problem or the physical meaning of the expected values, the P-spline definition is often modified by additional hypotheses, often translated into constraints on the solution or its derivatives. In this framework, our work is motivated by the aim of getting approximation models that fall within pre-established thresholds. Specifically, starting from a set of observed data, we consider a P-spline constrained between some prefixed bounds. In our paper, we just consider 0 as lower bound, although our approach applies to more general cases. We propose to get nonnegativity by imposing lower bounds on selected sample points. The spline can be computed through a sequence of linearly constrained problems. We suggest a strategy to dynamically select the sample points, to avoid extremely dense sampling, and therefore try to reduce as much as possible the computational burden. We show through some computational experiments the reliability of our approach and the accuracy of the results compared to some state-of-the-art models.

Keywords Regression splines · Constrained splines · P-splines · Algorithms for first-order approximation

Mathematics Subject Classification 65D07 · 65D10 · 65D15

✉ Rosanna Campagna
rosanna.campagna@unicampania.it

¹ Department of Mathematics and Physics, University of Campania “L. Vanvitelli”, Caserta, Italy

² Department of Environmental Sciences, Informatics and Statistics, Ca’ Foscari University of Venice, Venice, Italy

³ School of Mathematical Sciences, Dublin City University, Dublin, Ireland

1 Introduction

In the context of data analysis, splines are powerful and versatile nonparametric models, used for different purposes, e.g., building statistical distributions, investigating data structures, describing and predicting data behaviour [13, 20, 23].

Given a set of observed data (x_i, Y_i) , $i = 1, \dots, m$, with Y_i assumed to be some response of an unknown function $g : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$, corrupted by i.i.d. zero-mean Gaussian noise, i.e.,

$$Y_i = g(x_i) + v_i, \quad i = 1, \dots, m,$$

where $v_i \sim N(0, \sigma_g)$ and σ_g is the standard deviation of v_i , a cubic *regression spline* s is defined as a minimizer of the sum of the squared errors, that is

$$\min_{s \in \mathcal{S}} \sum_{i=1}^m (Y_i - s(x_i))^2, \quad (1.1)$$

or of a penalized least squares function

$$\min_{s \in \mathcal{S}} \sum_{i=1}^m \omega_i (Y_i - s(x_i))^2 + \lambda^2 \mathcal{P}(s), \quad (1.2)$$

with $\omega_1, \dots, \omega_m$ some given weights and $\lambda \neq 0$ regularization parameter. Here $\mathcal{S}$ is a spline space spanned by a suitable basis.

In (1.2) the residual sum of squares is balanced by a suitable penalty term $\mathcal{P}(s)$ expressed in terms of an integral or differential operator applied to s and/or its derivatives. In this last case, the spline is often called a *penalized regression spline*. The definition of the penalty characterizes the spline model. Different penalized regression splines are available in the literature, e.g. O-splines [16], P-splines [9], and T-splines [18]. They strongly differ from the classical *smoothing spline* (due to Whittaker [24] and then studied by Schoenberg [20], and for which Reinsch [17] proposed a well-known algorithm in 1967) which is the solution of the penalized least-squares problem

$$\operatorname{argmin}_{f \in W_2^p[a, b]} \sum_{i=1}^m (y_i - f(x_i))^2 + \lambda^2 \int_a^b (f^{(p)})^2 dx$$

in the Sobolev space

$$W_2^p[a, b] = \left\{ f : [a, b] \rightarrow \mathbb{R} : f, f', \dots, f^{(p-1)} \text{ absolutely cont., s.t. } \int_a^b (f^{(p)})^2 dx < \infty \right\}.$$

The main weakness of this model is the coincidence of nodes and smoothing data which, if numerous, as often happens in practice, can cause overfitting and model oscillations.

In contrast, penalized regression splines are defined by nodes possibly independent from the data, and this assumption makes them more flexible and suitable for data analysis.

We focus on P-splines [9], which are among the most successfully smoothers, used in a wide range of applications. A survey of the main developments on P-splines obtained in the last two decades is due to Eilers and Marx [10], who in 1996 introduced this family of splines with penalty defined through B-splines with equally-spaced knots [9]; the authors discard the derivative to express the roughness term as the sum of squares of differences of coefficients, making it extremely easy to compute.

P-splines inherit reproducing properties from B-splines and enjoy advantages of regression models, avoiding the problems of overfitting and consequent oscillations at the edges. In the applications, to face the different requirements imposed by the nature of the problem or the physical meaning of the expected values, the P-splines can be generalized to non-polynomial bases [4–6], or constrained by specific penalty terms forcing the model to be, e.g., monotone or positive. In [12] an application is described where the fitting by *positive* cubic splines is introduced to extract sparse and meaningful features from a set of nonnegative data vectors, while preserving their positive nature. In [11] the same authors prove that the functional form of a discrete dataset, through a linear combination of B-splines, enforces some intrinsic features on the recovered data, such as smoothness, particularly useful in inpainting or, more generally, image recovery problems. Since the B-splines are positive by nature [7], the positiveness is simply deduced by the projection of the coefficients in a nonnegative space. A conical combination of basis functions with a similar aim is also proposed in [25]. One must observe, however, that nonnegativity of the coefficients is only a sufficient condition to ensure the nonnegativity of a spline [8] and that it could be unnecessarily restrictive and result in an increase of the approximation error. The main goal of this work is to present a way to overcome such requirement for the construction of a nonnegative spline.

The design of shape constrained models, motivated by many applications, is achieved with different strategies; in [14, 19] the nonnegativity of P-splines is achieved through a data-driven node selection based on statistical techniques, while in [19] the spline fitting model is forced to be positive through a coefficient parametrization. Similar approaches can be found also in [21] with an algorithm for computing a smoothing polynomial spline with suitably prescribed constraints on the derivatives; semi-infinite constraints are replaced by finite ones, ending up with a least squares problem with linear inequality constraints. Finally, [22] presents a method that combines shape preservation and least squares approximation by splines with free knots, leading to a nonlinear least squares problem in both the coefficients and the knots.

This work is fundamentally motivated by the objective of getting approximation models that fall within pre-established thresholds. Specifically, starting from a set of observed data, we propose a constrained P-spline model aimed to build an approximating function bounded from some prefixed bounds. To facilitate the description we will focus only on the case of lower bounds providing some hints on how the method can be extended to the case of upper bounds on the spline. Moreover, in our work we consider 0 as lower bound, although the proposed approach applies to any bound, different bounds on different intervals.

We recall that there is extensive literature on numerical analysis and approximation theory that tackles the problem of providing upper bounds on several types of approximants or interpolants. This is usually achieved by deriving asymptotic bounds on the growth of the so-called Lebesgue constant of the approximation process, see e.g. [1–3]. We are however interested in bounds on the approximant both from below and from above, and especially we require that these bounds hold for any dataset, and not just in the asymptotic limit as the number of data grows to infinity. To this end, we introduce an algorithm and prove a bound that are both constructive and expressed in terms of the magnitude of the coefficients of the approximant.

B-splines are positive (bell-shaped) and compactly supported functions [7], and thus nonnegativity can be trivially imposed by considering only their combinations with nonnegative coefficients. However, as mentioned before this approach may prove too and unnecessarily restrictive. As an alternative we propose to introduce lower bounds on the solution of problem (1.2) on selected sample points that can be formulated as linear constraints on the objective function. This procedure may result reductive and unsatisfactory to guarantee the nonnegativity of the model since it ensures the nonnegativity only on the sample points. A strong limitation of this approach is that obtaining an acceptable result may require extremely dense sampling, thus involving a heavy computational burden. To overcome this drawback we propose a strategy in which the set of sampling points is updated dynamically. Our algorithm involves the alternation of two phases, one to add sampling points in critical areas (feeding), the other to eliminate unnecessary nodes from the set (pruning).

Similarly to the P-splines, higher orders models are straightforward to be deduced, since the higher order differences in the penalty term are easily defined.

A theoretical result is provided based on a suitable Lipschitz constraint condition that guarantees the positivity of our model, said CP-spline, over an entire interval $[a, b]$.

The paper is organized as follows: in Sect. 2 the constrained model is introduced and justified; the local nature is described, together with the motivation of dynamic local bounds. The algorithm is finally illustrated, equipped with procedures to prune the updating dataset and to drive the monotone behavior of the computed solution. A theoretical result about the global sign is also presented. Section. 3 presents numerical experiments that confirm the reliability of the model, the robustness of the algorithm, and the accuracy of the results compared to state-of-art models. Conclusions are drawn in Sect. 4.

2 CP-spline: a constrained P-spline model

In this section we first introduce P-splines, building our definition upon the assumptions and simplifications of the model introduced by [9]. Then, we define a CP-spline as a constrained P-spline model. Finally, we focus on a possible algorithm for adaptive selection of additional points to the data set, with the aim of forcing the model to be bounded from below in the data domain. For the sake of simplicity, only the problem of keeping the spline nonnegative will be considered hereinafter. We will refer to the

P-spline as the solution to problem (1.2), where the penalty is expressed in terms of the coefficients.

Definition 2.1 (P-Spline) Given a set of noisy data (x_i, Y_i) , $i = 1, \dots, m$, $a = x_1 < x_2 < \dots < b = x_m$, in $[a, b]$ and a cubic B-spline basis $\{B_j\}_{j=1, \dots, n}$ of dimension n , defined on the augmented set of $n + 6$ uniformly distributed knots, $\{\xi_{-1}, \dots, \xi_{n+2}\}$ with $\xi_2 \equiv x_1$ and $\xi_{n-1} \equiv x_m$, and introducing additional knots $\xi_{-1} < \xi_0 < \xi_1 < a$ and $b < \xi_n < \xi_{n+1} < \xi_{n+2}$, in order to span the spline space $\mathcal{S}$, we define the P-spline $s(x) \in \mathcal{S}$

$$s(x) = \sum_{j=1}^n a_j B_j(x), \quad (2.1)$$

as the spline whose coefficients are the solution of

$$\arg \min_{\mathbf{a} \in \mathbb{R}^n} \sum_{i=1}^m \omega_i (Y_i - s(x_i))^2 + \lambda^2 \sum_{j=3}^n \left((\Delta^2 \mathbf{a})_j \right), \quad (2.2)$$

with weights $\omega_1, \dots, \omega_m$, a regularization parameter λ , and where

$$(\Delta^2 \mathbf{a})_j = a_j - 2a_{j-1} + a_{j-2}, \quad j = 3, \dots, n.$$

If we set $\mathbf{y} \equiv (Y_1, \dots, Y_m)$, $\mathbf{W} = \text{diag}(\mathbf{w})$ with $\mathbf{w} \equiv (\omega_1, \dots, \omega_m)$, $\mathbf{V} = \sqrt{\mathbf{W}}$, $\mathbf{D} \in \mathbb{R}^{n \times n}$ as the tridiagonal matrix

$$\mathbf{D} = \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & \dots & \ddots & \dots & \vdots \\ 0 & \dots & 1 & -2 & 1 \\ 0 & \dots & \dots & 1 & -2 \end{pmatrix},$$

and $\mathbf{B} := (B_j(x_i))_{i=1, \dots, m}^{j=1, \dots, n}$ with band structure inherited by the B-splines locality, problem (2.2) can be written as

$$\arg \min_{\mathbf{a} \in \mathbb{R}^n} \|\mathbf{V}\mathbf{B}\mathbf{a} - \mathbf{V}\mathbf{y}\|^2 + \lambda^2 \mathbf{a}^T \mathbf{D}^T \mathbf{D} \mathbf{a} \Leftrightarrow \arg \min_{\mathbf{a} \in \mathbb{R}^n} \frac{1}{2} \mathbf{a}^T \left(\mathbf{B}^T \mathbf{W} \mathbf{B} + \lambda^2 \mathbf{D}^T \mathbf{D} \right) \mathbf{a} - \left(\mathbf{B}^T \mathbf{W} \mathbf{y} \right)^T \mathbf{a}. \quad (2.3)$$

The solution of the strictly convex quadratic problem (2.3) can be computed solving the linear system of normal equations

$$\left(\mathbf{B}^T \mathbf{W} \mathbf{B} + \lambda^2 \mathbf{D}^T \mathbf{D} \right) \mathbf{a} = \mathbf{B}^T \mathbf{W} \mathbf{y}. \quad (2.4)$$

Now we define the CP-spline and we focus on a possible adaptive selection of its defining sequence of points.

Definition 2.2 (CP-spline) Let s be a P-spline (see Definition 2.1), let furthermore $Z = \{z_1, z_2, \dots, z_p\} \subset [a, b]$ and $\mathcal{E} = \{\varepsilon_1, \dots, \varepsilon_p\} \subset [0, +\infty)$. Then a *constrained P-spline* (CP-spline), associated to the set of points Z (referred to as “sampling points”) and to the forcing sequence $\mathcal{E}$, is the function referred to as $s_{Z,\mathcal{E}} = \sum_{j=1}^n a_j B_j(x)$, whose coefficients solve the problem (2.2) with constraints:

$$\begin{aligned} \arg \min_{\mathbf{a} \in \mathbb{R}^n} & \sum_{i=1}^m \omega_i (Y_i - s(x_i))^2 + \lambda^2 \sum_{j=3}^n \left((\Delta^2 \mathbf{a})_j \right), \\ & \sum_{j=1}^n a_j B_j(z_i) \geq \varepsilon_i \geq 0, \quad i = 1, \dots, p. \end{aligned} \quad (2.5)$$

with weights $\omega_1, \dots, \omega_m$, a regularization parameter λ , and where

$$(\Delta^2 \mathbf{a})_j = a_j - 2a_{j-1} + a_{j-2}, \quad j = 3, \dots, n.$$

Because of (2.5), the function $s_{Z,\mathcal{E}}(x)$ is forced to be nonnegative in all points of the set Z . Actually, the ultimate goal of the constraints (2.5) is to keep $s_{Z,\mathcal{E}}(x)$ nonnegative over the entire interval $[a, b]$. A straightforward approach would suggest taking as Z grids of points equally spaced in the interval $[a, b]$, finer and finer until the result appears to be satisfactory according to some prefixed criterion. However a twofold drawback advises against such an approach. A very large number of points in Z , in the first place, makes the optimization problem computationally expensive. Secondly it could foster overfitting and oscillatory behaviour of the CP-spline. Therefore, the set Z should be reasonably small, with carefully chosen points able to favor, through the constraints (2.5), the nonnegativity of the P-spline. We propose to select a suitable set Z through an iterative algorithm. Given a current set Z (with a forcing sequence $\mathcal{E}$) the idea is to update it by (possibly) removing some points and adding new ones in those regions in which the spline is likely to become negative. This update must be followed by the forcing sequence update. Before describing our strategy, we consider the active set $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_p\}$, where the activity index $\alpha_i \in \{0, 1\}$ of z_i , $i = 1, \dots, p$, is defined as follows:

$$\alpha_i = \begin{cases} 1, & \text{if } s_{Z,\mathcal{E}}(z_i) = \varepsilon_i, \\ 0, & \text{otherwise.} \end{cases} \quad (2.6)$$

The kernel of the proposed algorithm, which we refer to as Algorithm 1, consists of the computation of the CP-spline and the management of the updating and/or pruning of the defined subsets Z , $\mathcal{E}$ and $\mathcal{A}$. More in details, Algorithm 2 outlines the strategy to update Z and $\mathcal{A}$ by adding new points at each iterate¹.

¹ In Algorithm 2 with the symbol $\cup$ we refer to *union and sorting*.

Algorithm 1 CP-spline

```

1: Input:
2:  $\Xi = \{\xi_{-1}, \dots, \xi_{n+2}\}$ , with  $\xi_i < \xi_{i+1}$ , uniform set of knots with  $\xi_2 \equiv a$ ,  $\xi_{n-1} \equiv b$ .
3:  $B_1(x), B_2(x), \dots, B_n(x)$  cubic B-splines defined on  $\Xi$ 
4:  $(x_i, Y_i), i = 1, \dots, m$  set of noisy data with  $x_i \in [a, b]$ 
5:  $Z = \{z_1, \dots, z_p\} \subseteq [a, b]$  with  $z_i < z_{i+1}$ ,  $\mathcal{A} = \{0, \dots, 0\}$ 
6: repeat
7:   compute the CP-spline  $s_{Z, \mathcal{E}}(x)$  solving (2.2) under constraints (2.5)
8:   update  $Z, \mathcal{A}$ 
9:   prune  $Z, \mathcal{A}$ 
10:  update  $\mathcal{E}$ 
11: until stop condition
12: Output
13:  $\bar{Z}, \bar{\mathcal{E}}, s_{\bar{Z}, \bar{\mathcal{E}}}(x)$ 

```

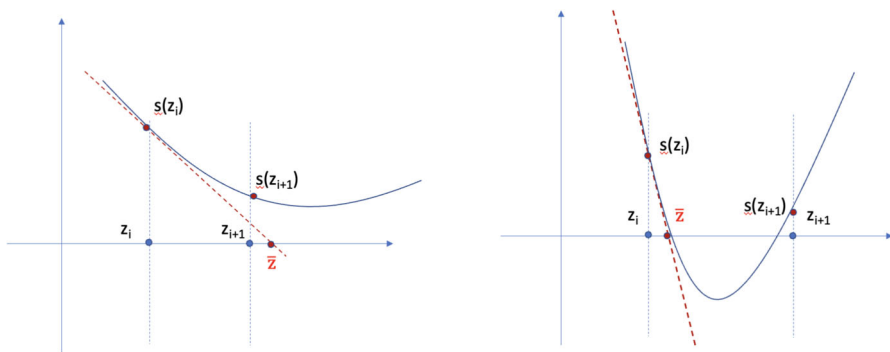


Fig. 1 Schematic illustration of the update step in Algorithm 2

The update algorithm adds points in those intervals in which the spline $s_{Z, \mathcal{E}}(x)$ is likely to be negative. This is done for each interval (z_i, z_{i+1}) in a twofold manner. If $\alpha_i = \alpha_{i+1} = 1$, then the spline may become negative in the interval (z_i, z_{i+1}) ; in this case, the midpoint of the interval is added to Z if the spline has negative derivative in z_i and positive in z_{i+1} . In the other case, for each interval (z_i, z_{i+1}) one computes the root $\bar{z}$ of the first order Taylor's expansion around one of the extreme points, and adds it to Z if $\bar{z} \in (z_i, z_{i+1})$ (see Fig. 1).

The algorithm may present some drawbacks, namely, adding too many (unnecessary) points in those regions in which the spline is close to the x-axis and fostering oscillatory behavior in such regions. About this last issue, we define $\mathcal{M} = \{j, \dots, j+q\}$, with $j \geq 1$ and $j+q \leq p$, *maximal active set* if $\alpha_i = 1$ for $i \in \mathcal{M}$ and $\alpha_i = 0$ for $i = j-1$ (if $j > 1$) and $i = j+q+1$ (if $j+q < p$). Moreover we denote with $\mathcal{J}^{\mathcal{M}}$ the interval $[x_{j-1}, x_{j+q+1}]$. We note that if $\mathcal{M}$ is a maximal active set, then the constraints (2.5) are active for $z_i \in \mathcal{J}^{\mathcal{M}}$; to smooth out the zigzagging behavior, we suggest to take for the corresponding ε_i small increasing [decreasing] values if the spline is likely to be increasing [decreasing]. To keep the number of points of Z under control we also fix a maximum number μ of points into any maximal active set $\mathcal{M}$. Our strategy is outlined in Algorithm 3. To further reduce

Algorithm 2 update $Z, \mathcal{A}$

```

1: Input:
2:  $Z = \{z_1, \dots, z_p\} \subseteq [a, b]$  with  $z_i < z_{i+1}$ ,  $\mathcal{E} = \{\varepsilon_1, \dots, \varepsilon_p\} \subseteq \mathbb{R}_0^+$ 
3:  $s_i = s_{Z, \mathcal{E}}(z_i)$ ,  $i = 1, \dots, m-1$ 
4:  $s'_i = s'_{Z, \mathcal{E}}(z_i)$ ,  $i = 1, \dots, m-1$ 
5:  $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_p\}$  where  $\alpha_i = \begin{cases} 1 & \text{if } s_i = \varepsilon_i \\ 0 & \text{otherwise} \end{cases}$ 
6:  $\bar{Z} = \emptyset$ ,  $\bar{\mathcal{A}} = \emptyset$ 
7: for  $i = 1, \dots, m-1$  do
8:   if  $s_i = \varepsilon_i$  and  $s_{i+1} = \varepsilon_{i+1}$  then
9:      $\bar{z} = (z_i + z_{i+1})/2$ 
10:    if  $s'_i < 0$  and  $s'_{i+1} > 0$  then  $\bar{Z} = \bar{Z} \cup \{\bar{z}\}$ ;  $\bar{\mathcal{A}} = \bar{\mathcal{A}} \cup \{1\}$ 
11:    end if
12:  else if  $s_i > s_{i+1}$  and  $s_i > \varepsilon_i$  then
13:     $\bar{z} = z_i - s_i/s'_i$ 
14:    if  $\bar{z} \in (z_i, z_{i+1})$  then  $\bar{Z} = \bar{Z} \cup \{\bar{z}\}$ ;  $\bar{\mathcal{A}} = \bar{\mathcal{A}} \cup \{1\}$ 
15:    end if
16:  else if  $s_i < s_{i+1}$  and  $s_{i+1} > \varepsilon_{i+1}$  then
17:     $\bar{z} = z_{i+1} - s_{i+1}/s'_{i+1}$ 
18:    if  $\bar{z} \in (z_i, z_{i+1})$  then  $\bar{Z} = \bar{Z} \cup \{\bar{z}\}$ ;  $\bar{\mathcal{A}} = \bar{\mathcal{A}} \cup \{1\}$ 
19:    end if
20:  end if
21: end for
22: Output
23:  $Z = Z \cup \bar{Z} = \{z_1, \dots, z_q\}$  with  $z_i < z_{i+1}$ ;  $\mathcal{A} = \mathcal{A} \cup \bar{\mathcal{A}} = \{\alpha_1, \dots, \alpha_q\}$ 

```

the number of elements in Z the pruning strategy is combined with a simple dropping one, based on the magnitude of $s_{Z, \mathcal{E}}(z_i)$.

Algorithm 3 prune $Z, \mathcal{A}$

```

1: Input:  $Z = \{z_1, \dots, z_p\} \subseteq [a, b]$  with  $z_i < z_{i+1}$ ,  $\mathcal{E} = \{\varepsilon_1, \dots, \varepsilon_p\} \subseteq \mathbb{R}_0^+$ ,  

 $\mathcal{A} = \{\alpha_1, \dots, \alpha_p\}$ ,  $\alpha_i \in \{0, 1\}$ ,  $\mu \in \mathbb{N}$ ,  $\bar{\varepsilon} \in \mathbb{R}^+$ 
2: for  $i = 1, \dots, m-1$  do
3:    $s_i = s_{Z, \mathcal{E}}(z_i)$ 
4:   if  $s_i > \bar{\varepsilon}$  then
5:      $Z = Z \setminus \{z_i\}$ 
6:   end if
7: end for
8: for  $\mathcal{M} = \{j, \dots, j+q\}$ , with  $j \geq 1$  and  $j+q \leq p$  do
9:   if  $q > \mu$  then
10:    replace the points  $z_j, z_{j+1}, \dots, z_{j+q}$  with  $z_j + kh$ ,  $k = 1, \dots, \mu$ ,  $h = (z_{j+q} - z_j)/\mu$ 
11:    update consequentially  $\mathcal{A}$  with activity index  $\alpha_i = 1$  for the new points
12:   end if
13: end for
14: Output: updated  $Z$  and  $\mathcal{A}$ 

```

Summing up, the algorithm computes a finite sequence of CP-splines w.r.t. to a set Z that is gradually modified by adding and removing points to foster the spline nonnegativity. The *update* steps aim to locate areas in which the spline tends to be

Algorithm 4 update $\mathcal{E}$

```

1: Parameters:  $\epsilon$ 
2: Input:  $\mathcal{E} = \{\varepsilon_1, \dots, \varepsilon_p\} \subseteq \mathbb{R}_0^+$ ,  $\mathcal{A} = \{\alpha_1, \dots, \alpha_p\}$ ,  $\alpha_i \in \{0, 1\}$ 
3:  $s_i = s_{Z, \mathcal{E}}(z_i)$ ,  $i = 1, \dots, m-1$ 
4: for  $\mathcal{M} = \{j, \dots, j+q\}$ ,  $j \geq 1$ ,  $j+q \leq p$  do
5:   if  $s_{j-1} > s_{j+q+1}$  then
6:     set  $\varepsilon_{i+i} = (i+1)\epsilon$ ,  $i = 0, \dots, q$ 
7:   else
8:     set  $\varepsilon_{i+i} = (q-i+1)\epsilon$ ,  $i = 0, \dots, q$ 
9:   end if
10: end for
11: Output: updated  $\mathcal{E}$ 

```

negative, whereas the *prune* steps avoid over-concentration of points in such areas. We note that the sequence of minimization problems of the form (2.2) share the same objective function (i.e., (2.2)) and differ only in the constraints (2.5). This allows one to reduce the computational cost by using the solution of the previous problem to warmstart the optimization.

Remark 2.1 Though the algorithm was presented for the case of nonnegatively constrained splines (i.e., lower bound equal to 0), it can be easily extended to the case of generic lower and upper bounds. In detail, when it comes to upper bounds, it is sufficient to observe that one can introduce a set of sampling points $Z^u = \{z_1^u, z_2^u, \dots, z_k^u\} \subset [a, b]$ for the upper bounds resulting in the introduction into (2.5) of inequality constraints of the form

$$\sum_{j=1}^n a_j B_j(z_i^u) \leq U_i, \quad i = 1, \dots, k,$$

for given upper bounds U_i . The set of upper-bound sampling points Z^u can then be updated by using procedures analogous to the ones described in Algorithms 2–4.

We conclude this section by stating a constructive condition that can be used to verify a-posteriori that the computed CP-spline is positive.

Theorem 2.1 *Under the assumptions of Definition 2.2, let $s_{Z, \mathcal{E}}$ be the CP-spline associated to Z and $\mathcal{E}$. Define $z_0 := a$, $z_{p+1} := b$, so that $[a, b] = \cup_{i=0}^p [z_i, z_{i+1}]$. For $i \in \{0, \dots, p\}$ set $h_i := z_{i+1} - z_i$ and*

$$\bar{\varepsilon}_i := \begin{cases} \max(\varepsilon_i, \varepsilon_{i+1}), & i \in \{1, \dots, p-1\}, \\ \varepsilon_1, & i = 0, \\ \varepsilon_p, & i = p+1, \end{cases}$$

and let $L_i := L(s_{Z, \mathcal{E}}, [z_i, z_{i+1}])$ be the Lipschitz constant of $s_{Z, \mathcal{E}}$ in $[z_i, z_{i+1}]$. Then for each $i \in \{0, \dots, p\}$, if

$$\bar{\varepsilon}_i > L_i h_i, \tag{2.7}$$

then $s_{Z,\mathcal{E}}(x) > 0$ for all $x \in [z_i, z_{i+1}]$.

Proof Let $i \in \{0, \dots, p\}$. By definition of L_i and h_i , for all $z_i \in Z$ and $z \in [z_i, z_{i+1}]$ we have

$$s_i - s(z) = s(z_i) - s(z) \leq |s(z) - s(z_i)| \leq L_i |z - z_i| \leq L_i h_i,$$

and by rearranging this inequality we get

$$s(z) \geq s_i - L_i h_i \quad \forall z \in [z_i, z_{i+1}],$$

and by the same argument, if $i \in \{1, \dots, p\}$, we have

$$s(z) \geq s_{i+1} - L_i h_i \quad \forall z \in [z_i, z_{i+1}].$$

It follows that for if $i \in \{1, \dots, p\}$ we have

$$\begin{aligned} s(z) &\geq \max(s_i - L_i h_i, s_{i+1} - L_i h_i) = \max(s_i, s_{i+1}) - L_i h_i = \max(\varepsilon_i, \varepsilon_{i+1}) - L_i h_i \\ &= \bar{\varepsilon}_i - L_i h_i, \quad \forall z \in [z_i, z_{i+1}], \end{aligned}$$

while if $i \in \{0, p+1\}$ we have

$$\begin{aligned} s(z) &\geq s_1 - L_1 = \bar{\varepsilon}_1 - L_1 h_1, \quad \forall z \in [z_i, z_{i+1}] \\ s(z) &\geq s_p - L_p = \bar{\varepsilon}_p - L_p h_p, \quad \forall z \in [z_p, z_{p+1}]. \end{aligned}$$

Thus for all $i \in \{0, \dots, p+1\}$ the condition $\bar{\varepsilon}_i > L_i h_i$ implies $s(z) > 0$ for $z \in [z_i, z_{i+1}]$, which proves the statement. $\square$

Remark 2.2 Apart from this kind of a-posteriori verifications, we remark that the algorithm is based on a heuristic argument and that there are corner cases where a spline passes the controls of Algorithm 2 but has negative values. A theoretical guarantee on the non-negativity of the spline would thus need to rule out these cases, which are difficult to control due to the complex interaction between the discretization, the thresholds, and the regularization parameter.

However, given the controls of Algorithm 2 it should be expected that these corner cases would need to have large values and variations of the second derivative. This behavior should be prevented by a large enough regularization parameter λ .

3 Numerical experiments

In this section, we provide some numerical tests to evaluate the practical behaviour of the proposed algorithm compared to other state-of-the-art methods, on some synthetic test problems and a real-life test problem based on COVID-19 infection data. All experiments were performed on a single laptop equipped with a Intel(R) Core(TM) Ultra 5 125U CPU @ 3.60GHz with 16 GB RAM, using MATLAB R2024a and Python 3.12.

We performed three sets of experiments with different purposes: first, we compared our CP-spline method with the unconstrained P-spline and the P-spline with nonnegative coefficients, here denoted as NNP-spline; then, we evaluated the difference between the performances obtained by CP-spline algorithm and the method proposed in [12], where a different approach to perform projection onto nonnegative set is developed; finally, we performed experiments on both synthetic and real-life datasets to compare the performance of the Python implementation of the proposed method with the method presented in [15].

In our experiments we set $\mathbf{W} = \mathbf{I}_m$, identity matrix, in (2.3); moreover, for the sake of the practical implementation, we note that problem (2.2) with constraint (2.5) can be reformulated as follows

$$\begin{aligned} \operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^n} \quad & \frac{1}{2} \mathbf{a}^T \mathbf{H} \mathbf{a} + \mathbf{a}^T \mathbf{f}, \\ \text{s.t.} \quad & \mathbf{B} \mathbf{a} \geq \boldsymbol{\varepsilon} \end{aligned} \quad (3.1)$$

with $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)$, $\mathbf{H}$ and $\mathbf{f}$ suitably defined by exploiting $\mathbf{B}$ and $\mathbf{D}$ in (2.4). Indeed, the objective function in (2.2) becomes

$$\mathbf{a}^T \tilde{\mathbf{B}} \mathbf{a} - 2\mathbf{a}^T \tilde{\mathbf{b}} + \lambda^2 \mathbf{a}^T \mathbf{T} \mathbf{a},$$

where

$$\begin{aligned} b^{(i)} &= (B_j(x_i))_{j=1, \dots, n} \in \mathbb{R}^{n \times 1}, \\ \tilde{\mathbf{B}}^{(i)} &= b^{(i)} b^{(i)T}, \text{ i.e., } (\tilde{\mathbf{B}}^{(i)})_{h,k} = B_h(x_i) B_k(x_i), \end{aligned}$$

and denoting

$$\tilde{\mathbf{b}} = \sum_i y_i b^{(i)} \in \mathbb{R}^{n \times 1}$$

Moreover, by setting

$$t_j(k) = \begin{cases} 1, & k = j - 2, \\ -2, & k = j - 1, \\ 1, & k = j, \\ 0, & \text{otherwise,} \end{cases} \quad (t_j \in \mathbb{R}^{n \times 1}) \quad (3.2)$$

we define

$$\mathbf{T} = \sum_3^n t_j t_j^T \in \mathbb{R}^{n \times n}.$$

Finally, with

$$\mathbf{Q} = \sum_i \tilde{\mathbf{B}}^{(i)} \in \mathbb{R}^{n \times n}$$

we write $\mathbf{H} = \mathbf{Q} + \lambda^2 \mathbf{T}$ and $\mathbf{f} = -2\tilde{\mathbf{b}}$, thus obtaining formulation (3.1).

The first sets of experiments were performed on MATLAB 2024a. In our MATLAB implementation of Algorithm 1, problem (3.1) is solved by using the `lsqlin` function, which allows one to solve linear least-squares with linear inequality constraints through an active-set method.

The regularization parameter λ in $\mathbf{H}$ is computed by L-curve criterion. The Generalized Cross Validation (GCV) algorithm has also been tested with similar results.

We considered five test problems, named TP1, TP2, TP3, TP4 and TP5, defined as follows:

- TP1: $n = 15$ knots linearly spaced in $[-20, 20]$, $m = 50$ data points $(x_i, \tilde{Y}_i)_{i=1, \dots, m}$ where, for any $i = 1, \dots, m$, x_i are randomly distributed, $\tilde{Y}_i = f(x_i) + v_i$ with $f(x) = \frac{1}{1.5\sqrt{2\pi}} e^{-\frac{1}{2} \frac{x^2}{1.5^2}}$, and the perturbation v_i is obtained from a normal distribution with zero mean and standard deviation $1.0e - 3$.
- TP2: $n = 15$ knots linearly spaced in $[1, 3]$, $m = 30$ data points $(x_i, \tilde{Y}_i)_{i=1, \dots, m}$ where, for any $i = 1, \dots, m$, x_i are linearly spaced, $\tilde{Y}_i = f(x_i) + v_i$ with $f(x) = \frac{-1}{e^{\sin(6x)^2}}$, and the perturbation v_i is obtained from a normal distribution with zero mean and standard deviation 0.015 .
- TP3: $n = 9$ knots linearly spaced in $[0, 10]$, $m = 100$ data points $(x_i, Y_i)_{i=1, \dots, m}$ where, for any $i = 1, \dots, m$, x_i are linearly spaced, and $Y_i = f(x_i)$ are obtained from a cubic spline interpolation.
- TP4: $n = 9$ knots linearly spaced in $[0, 2\pi]$, $m = 100$ data points $(x_i, Y_i)_{i=1, \dots, m}$ where, for any $i = 1, \dots, m$, x_i are linearly spaced, and $Y_i = f(x_i)$ with $f(x) = \sin(x) + 0.9$.
- TP5: $n = 9$ knots linearly spaced in $[0, 5]$, $m = 100$ data points $(x_i, Y_i)_{i=1, \dots, m}$ where, for any $i = 1, \dots, m$, x_i are linearly spaced, and $Y_i = f(x_i)$ with $f(x) = e^{-x} \cos(x)$.

The main loop in Algorithm 1 is stopped when the computed solution is nonnegative within a maximum number of 20 iterations. Moreover, a safeguarding condition is included to control the number of points in Z to avoid an excessive increasing of its size. In Algorithm 3 we set $\mu = 5$ and $\bar{\epsilon} = \mu \cdot \epsilon$ whereas in Algorithm 4, $\epsilon = 10^{-8}$.

The results obtained on these test problems are reported in Figs. 1, 2, 3 and 4. In Fig. 3, we evaluated the ability of CP-spline algorithm to approximate a perturbed Gaussian distribution (TP1) and data points generated as perturbation of an oscillating positive function (TP2), comparing the results with those obtained by using the basic P-spline regression and P-spline with nonnegative coefficients. From these plots, we can observe how the CP-spline algorithm can better approach the peaks, resulting in a more faithful approximation of the original function (starting from few data points), compared to the other methods. Also, CP-spline seems to better capture the extreme

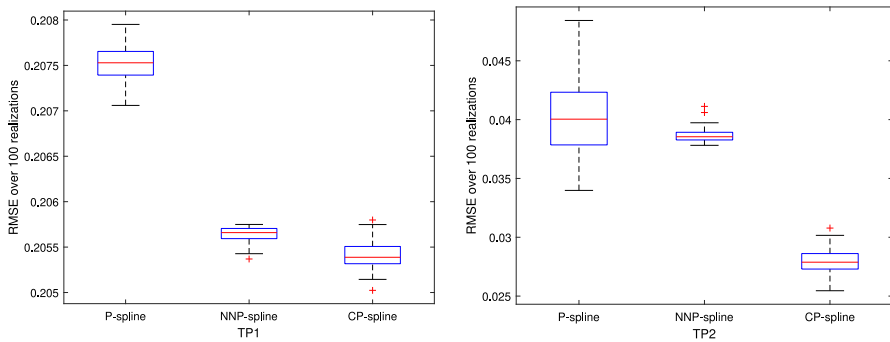


Fig. 2 Boxplot distributions of the RMSE values over 100 random realizations of the error on the data for TP1 (left panel) and TP2 (right panel)

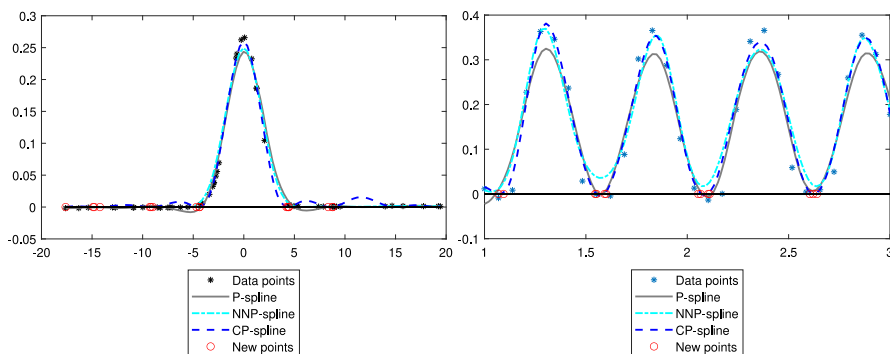


Fig. 3 Numerical results on synthetic test problem TP1 (left) and TP2 (right)

values in the nonnegative regions. This is confirmed by the RMSE values, which are given in Table 1. In Fig. 3, we also show the effect of the strategy on test problems TP1 and TP2, by plotting the new data points (red circles).

Figure 4 shows a comparison between our CP-spline method and the projection method proposed in [12]. These plots highlight the local nature of the CP-spline method that, indeed, can foster then nonnegativity of the function f while preserving the original graph in the feasible points where the constraints have no effect; in contrast, the projection method derived from [12] seems to cause a loss of the original information, thus producing a global action on the function f . This visual insight is confirmed also by the root mean squared errors (RMSE) reported in Table 2.

The effectiveness and convergence of the procedure have been tested through repeated experiments for TP1 and TP2, to assess the performance of the algorithm over 100 realizations of the error on the data points (see Fig. 2).

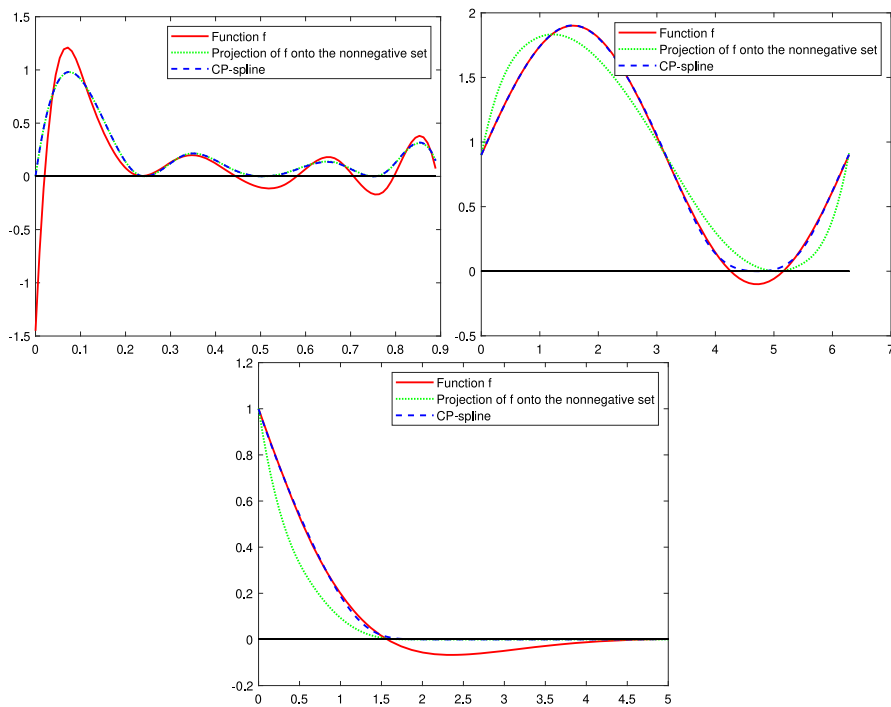


Fig. 4 Numerical results on synthetic test problem TP3 (top left), TP4 (top right), and TP5 (bottom)

3.1 Comparison with other software

To conclude the numerical experiments section we present a comparison we performed on the `cpsplines` library², based on the work in [15]. The comparison is relevant since the `cpsplines` library exploits a reformulation of the problem to a second-order cone program leading to theoretical nonnegativity guarantees.

Since the library was written in Python language, to ensure a fair comparison in terms of computational times, we built our own Python implementation of the algorithm presented in this work. Before presenting the results of this test it is important to highlight some details in the Python implementation:

- for the selection of the optimal regularization parameter we made use of the Python `lcurve` library³;
- to use the `lcurve` method, a GSVD must be computed first; since a function for GSVD was not available neither in NumPy nor in SciPy linear algebra modules, we created our own implementation of the GSVD, following the same algorithm used in MATLAB's `gsvd` function;

² <https://github.com/ManuelNavarroGarcia/cpsplines>.

³ <https://github.com/eric-brandao/lcurve>.

Table 1 Root mean squared errors (RMSE) computed on the m data points of test problems TP1 and TP2

	P-spline	NNP-spline	CP-spline
TP1	$4.62 \cdot 10^{-4}$	$5.04 \cdot 10^{-4}$	$1.02 \cdot 10^{-4}$
TP2	$8.2 \cdot 10^{-3}$	$8.68 \cdot 10^{-3}$	$4.39 \cdot 10^{-3}$

Table 2 Root mean squared errors (RMSE) computed on the m data points of test problems TP3, TP4 and TP5

	Projection [12]	CP-spline
TP3	$8.56 \cdot 10^{-2}$	$8.56 \cdot 10^{-2}$
TP4	$1.33 \cdot 10^{-2}$	$1.16 \cdot 10^{-2}$
TP5	$7.57 \cdot 10^{-2}$	$4.69 \cdot 10^{-3}$

- to solve the linearly-constrained least-squares problems arising at each iteration of the proposed algorithm we resorted to the `quadprog` library⁴.

For this comparison we considered the five problems used in the previous experiments to which we added two problems derived from the `cpsplines` library:

- GAUSS: approximation of simulated data affected by Gaussian noise of the form

$$y_i = f(x_i) + v_i,$$

where $x_i = 0, 1, 2, \dots, 200$,

$$f(x) = e^{(4 - \frac{x}{25})} + 4 \cos\left(\frac{x}{8}\right),$$

and the perturbation v_i is obtained from a normal distribution with zero mean and standard deviation 3

- COVID: smoothing of the normalized number of infections in Comunitat Valenciana over 586 days starting from the first event date to 4th October 2021.

We ran the two methods on each instance for 100 times and measured mean and standard deviation of the execution times. Additionally, we measured the relative distance in ℓ_∞ norm between the solution computed by the two methods. We report all the results in Table 3 and in Fig. 5 the approximations obtained by the two methods for problems TP1, TP5, GAUSS, and COVID.

From the table, one can clearly observe that the proposed method in general presents lower or comparable execution times than the ones from the `cpsplines` and has always a smaller standard deviation. From the table one can clearly see that the computational performance difference between the two algorithms is particularly evident in two cases: TP5, in which `cpsplines` is 3 times faster, and COVID, in which our method is 5 times faster. This doesn't come completely unexpected. Indeed, if one observes the two problems one can see that they correspond to two extreme opposite scenarios (see Fig. 5).

⁴ <https://github.com/quadprog/quadprog>.

Table 3 Mean and standard deviation of the execution times for the two Python-based codes with distance between the solutions

	CP-spline		cpsplines library		ℓ_∞ -distance
	Mean time (s)	Std	Mean time (s)	Std	
TP1	0.065	<i>0.021</i>	0.103	<i>0.021</i>	$1.52 \cdot 10^{-1}$
TP2	0.103	<i>0.017</i>	0.190	0.038	$3.39 \cdot 10^{-2}$
TP3	0.112	<i>0.017</i>	0.117	0.046	$4.49 \cdot 10^{-3}$
TP4	0.155	<i>0.017</i>	0.157	0.044	$8.55 \cdot 10^{-3}$
TP5	0.346	<i>0.034</i>	0.154	0.044	$8.73 \cdot 10^{-3}$
GAUSS	0.329	<i>0.035</i>	0.325	0.055	$4.49 \cdot 10^{-3}$
COVID	1.069	<i>0.084</i>	6.397	1.23	$1.13 \cdot 10^{-2}$

Best time and smallest standard deviation have been highlighted in bold and italic fonts, respectively

The data in problem TP5 present a long tail approaching the x-axis. This forces the proposed method to run for 12 iterations to identify 6 well-placed sample point leading to the resulting nonnegative spline. Problem COVID has sufficiently positive data and the only region in which control is needed is at the origin, where our method had to run 3 iterations to identify 3 sample points. It is worth recalling that our method solves a linearly-constrained quadratic programming problem at each iteration and the number of constraints at the beginning is equal to the number of data points (since we use data points to initialize the vector of sample points). On the contrary, the `cpsplines` library solves a single second-order cone program through the MOSEK⁵ interior-point solver, which, however, can become very expensive when the problem size increases.

Another peculiar case from Table 3 is TP1, where the distance between the two approximations is two orders of magnitude higher with respect to other cases. It is interesting to observe the two curves to understand why it is like this. This can be done by looking at the top-left tile of Fig. 5. In fact, one can observe that the approximation built from the proposed method better approximates the original data, which is a symptom of the potential practical benefits of the “local” constraints in the proposed method over the global nonnegativity constraint imposed in the `cpsplines` formulation.

Overall, the results suggest that, despite the lack of a theoretical guarantee of positiveness of the solution, the proposed heuristic strategy can compute splines which compare well qualitatively with the ones from `cpsplines` and require a lower computational effort.

4 Conclusions

In our article, we addressed the problem of designing smooth curves with specific requirements on the variability range in the general framework of data approximation

⁵ <https://www.mosek.com/>.

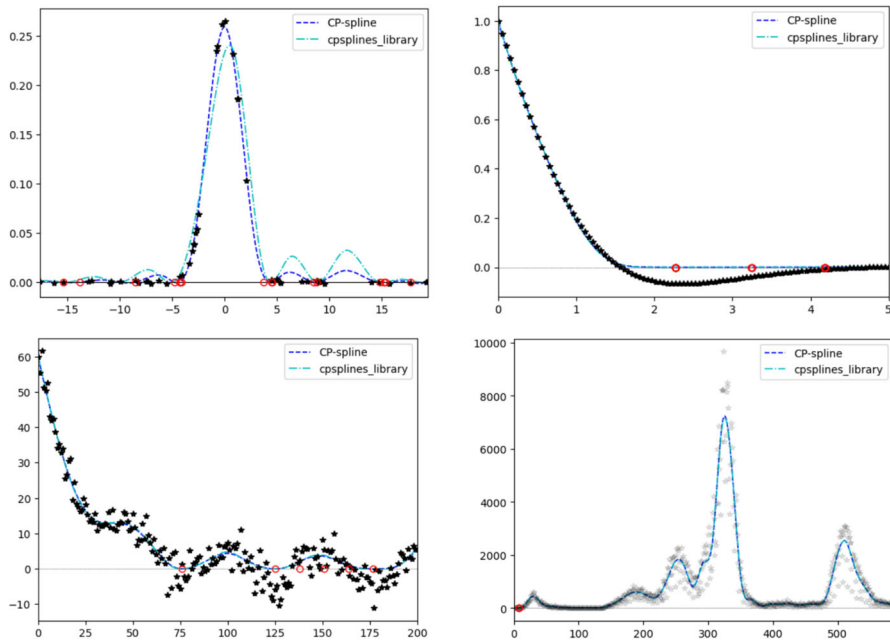


Fig. 5 Non-negative splines obtained from the Python libraries for problems TP1 (top left), TP5 (top right), GAUSS (bottom left) and COVID (bottom right). Red circles represent sample points for the CP-spline approximation

models. An algorithm is proposed to construct a constrained P-spline model, where such requirements are forced to be satisfied by explicitly imposing model-bound constraints on a set of dynamically updated sampling points. The update heuristic strategy is based on a first-order local approximation. Appreciable numerical results seem to encourage further investigations and developments of our algorithm. From a theoretical point of view, in the direction of finding valid results regarding the (possibly finite) convergence of the method. From a practical point of view, developing ad-hoc optimization algorithms for the solution of the least-squares minimization problems at each iteration of the proposed algorithm to reduce the computational cost of the method in cases in which a large number of iterations is required.

Acknowledgements The authors RC, SC, GS, and GT are members of the INdAM research group GNCS, which has partially supported this work. This research has been accomplished within RITA (Research Italian network on Approximation) and UMI-TAA groups, of which the corresponding author is a member. This work has been partially supported also by the project COMETA - Computational METHODS in data analysis (CUP B63C23000650005) of University of Campania “Luigi Vanvitelli”. We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on February 02, 2022 by the Italian Ministry of University and Research (MUR), funded by the European Union - NextGenerationEU- Project Title “Numerical Optimization with Adaptive Accuracy and Applications to Machine Learning” - CUP E53D23007700006 - Grant Assignment Decree No. 764 adopted on June 05, 2023 by the Italian Ministry of University and Research (MUR). We also acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 1409 published on September 14, 2022 by the Italian Ministry of University and Research (MUR), funded by the European Union - NextGenerationEU-

Project Title “A multidisciplinary approach to evaluate ecosystems resilience under climate change” - CUP B53D23027910001 - Grant Assignment Decree No.1379, adopted on September 01, 2023 by the Italian Ministry of University and Research (MUR). The authors thank the anonymous referees for their helpful comments and suggestions that improved the quality of the manuscript.

Funding Open access funding provided by Università degli Studi della Campania Luigi Vanvitelli within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Bandiziol, C., De Marchi, S.: On the Lebesgue constant of the trigonometric Floater-Hormann rational interpolant at equally spaced nodes, *Dolomites Research Notes Approx.*, 12 (2019), <https://doi.org/10.14658/PUPJ-DRNA-2019-1-6>
2. Bos, L., Marchi, S.D., Hormann, K., Klein, G.: On the Lebesgue constant of barycentric rational interpolation at equidistant nodes. *Numer. Math.* **121**, 461–471 (2012). <https://doi.org/10.1007/s00211-011-0442-8>
3. Bos, L., Marchi, S.D., Hormann, K., Sidon, J.: Bounding the Lebesgue constant of Berrut's rational interpolant at general nodes. *J. Approx. Theory* **169**, 7–22 (2013). <https://doi.org/10.1016/j.jat.2013.01.004>
4. Campagna, R., Conti, C.: Penalized hyperbolic-polynomial splines. *Applied Mathematics Letters* **118**, 107159 (2021). <https://doi.org/10.1016/j.aml.2021.107159>
5. Campagna, R., Conti, C.: Reproduction capabilities of penalized hyperbolic-polynomial splines, *Applied Mathematics Letters*, 132, 108133, ISSN 0893-9659 (2022), <https://doi.org/10.1016/j.aml.2022.108133>
6. Campagna, R., Conti, C., Cuomo, S.: A linear algebra approach to hp-splines frequency parameter selection. *Applied Mathematics and Computation* **458**, 128241 (2023). <https://doi.org/10.1016/j.amc.2023.128241>
7. de Boor, C.: *A Practical Guide to Splines*, Applied Mathematical Sciences, vol. 27, Springer, New York (1978), <https://doi.org/10.1007/978-1-4612-6333-3>
8. de Boor, C., Daniel, J.W.: Splines with nonnegative b-spline coefficients. *Mathematics of Computation* **28**, 565–568 (1974). <https://doi.org/10.1090/S0025-5718-1974-0378357-2>
9. Eilers, P.H.C., Marx, B.D.: Flexible smoothing with B-splines and penalties. *Statistical Science* **11**(2), 89–121 (1996). <https://doi.org/10.1214/SS/1038425655>
10. Eilers, P.H.C., Marx, B.D., Durbán, M.: Twenty years of p-splines, *Sort-statistics and Operations Research Transactions*, 39, 149–186 (2015), <https://api.semanticscholar.org/CorpusID:12430996>
11. Hautecoeur, C., Glineur, F.: Image completion via nonnegative matrix factorization using HALS and B-splines. In: *ESANN 2020, 28th European Symposium on Artificial Neural Networks - Computational Intelligence and Machine Learning* (Bruges, Belgium, du 02/10/2020 au 04/10/2020) (2020), <http://hdl.handle.net/2078.1/229855>
12. Hautecoeur, C., Glineur, F.: Nonnegative matrix factorization over continuous signals using parametrizable functions. *Neurocomputing* **416**, 256–265 (2020). <https://doi.org/10.1016/j.neucom.2019.11.109>

13. Lyche, T., Manni, C., Speleers, H.: Foundations of Spline Theory: B-Splines, Spline Approximation, and Hierarchical Refinement, pp. 1–76, Springer, Cham (2018), https://doi.org/10.1007/978-3-319-94911-6_1
14. Maturana-Russel, P., Meyer, R.: Bayesian spectral density estimation using p-splines with quantile-based knot placement. *Computational Statistics* **36**, 2055–2077 (2021). <https://doi.org/10.1007/s00180-021-01066-7>
15. Navarro-García, M., Guerrero, V., Durban, M.: On constrained smoothing and out-of-range prediction using P-splines: A conic optimization approach. *Applied Mathematics and Computation* **441**, 127679 (2023). <https://doi.org/10.1016/j.amc.2022.127679>
16. O’Sullivan, F.: A statistical perspective on ill-posed inverse problems. *Statistical Science* **1**(4), 502–518 (1986). <https://doi.org/10.1214/ss/1177013525>
17. Reinsch, C.: Smoothing by spline functions. *Numer. Math.* **10**(10), 177–183 (1967). <https://doi.org/10.1007/BF02162161>
18. Ruppert, D., Carroll, R.J.: Theory & methods: Spatially-adaptive penalties for spline fitting. *Australian & New Zealand Journal of Statistics* **42**(2), 205–223 (2000). <https://doi.org/10.1111/1467-842X.00119>
19. Schellhase, C., Kauermann, G.: Density estimation and comparison with a penalized mixture approach. *Comput Stat* **27**, 757–777 (2012). <https://doi.org/10.1007/s00180-011-0289-6>
20. Schoenberg, I.J.: Spline functions and the problem of graduation. *Proc. Nat. Acad. Sci.* **52**(52), 947–950 (1964)
21. Schwetlick, H., Kunert, V.: Spline smoothing under constraints on derivatives. *BIT Numerical Mathematics* **33**, 512–528 (1993). <https://doi.org/10.1007/BF01990532>
22. Schütze, T., Schwetlick, H.: Constrained approximation by splines with free knots. *BIT Numerical Mathematics* **37**, 105–137 (1997). <https://doi.org/10.1007/BF02510176>
23. Wang, Y.: *Smoothing Splines: Methods and Applications* (1st ed.), p. 384, Chapman and Hall/CRC, London (2011), <https://doi.org/10.1201/b10954>
24. Whittaker, E.T.: On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society* **41**, 63–75 (1922). <https://doi.org/10.1017/S0013091500077853>
25. Zdunek, R., Cichocki, A., Yokota, T.: B-spline smoothing of feature vectors in nonnegative matrix factorization, In: *International Conference on Artificial Intelligence and Soft Computing*, pp. 72–81 (2014), https://doi.org/10.1007/978-3-319-07176-3_7

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.