Eawag
Swiss Federal Institute of Aquatic Science and Technology

Spike
Sensor management, Process Intelligence and Control at Eawag

Technical Report No. 2, v1.0

(TR-002-01-0)

Bounding Procedures for Fitting of Shape Constrained Splines to Multivariate Data Series

Author:
Kris Villez

Dübendorf, Switzerland
Originally published: 19/08/2015
Last update: 19/08/2015

K. Villez (2015). *Bounding Procedures for Fitting of Shape Constrained Splines to Multivariate Data Series.* Eawag, Dübendorf, Switzerland, Technical Report TR-002-01-0.

Abstract

Previous work on qualitative trend analysis (QTA) has led to the development of a shape constrained splines (SCS) method. This method delivers a deterministic and globally optimal solution for the location of so called transitions in a qualitative representation (QR). Unfortunately, the method is limited to the analysis of univariate data series. This technical report describes and proves the bounding procedures that are necessary to expand the method to multivariate data series.

1 Introduction

For the purpose of shape constrained function fitting it is important to distinguish between several ways multivariate properties of data can be interpreted. The following three are identified:

1. Measurements are independent repetitions of each other
2. Measurements are different in value yet reflect the same underlying state of the measured system
3. Measurements are repeated across spatial or time dimensions, reflecting meaningful changes in the measured system. Across multiple spatial and/or time dimensions, these changes can be additive or non-additive.

In the a following section, the bounding procedures are given for multivariate data considering the first and second way. Bounding procedures accounting for the third way are not available yet.

2 Definitions and notation

The following definitions are used in this work:

Episode. An episode is an argument interval over which the signs of al set of functions and/or data series and/or one or more of their derivatives are the same and do not change. An episode is defined by a primitive, a start time, and an end time.

Primitive. A primitive is a unique combination of signs for a value of one or more functions and/or one or more of their derivatives. Each primitive is usually referred to by means of an arbitrary chosen character. In this work, the sign of the first and second derivatives are of interest. The correspondence between the signs of the derivatives and the characters used in this study are given in Fig. 1.

Qualitative sequence. A qualitative sequence (QS) is a series of primitives. Such a QS is used to describe the assessed or expected shape of one or more functions or multivariate time series. A QS does not include the argument locations (transitions) at which a change in primitive is expected or observed.

Transition. A transition is defined as the argument location of a change-point between two sequential primitives.

Qualitative representation. A qualitative representation is a complete description of the expected or observed shape of a function or time series and consists of a QS and values for the argument values of the corresponding transitions and is equivalent to a series of contiguous episodes.

A qualitative sequence is defined mathematically by means of symbols, $s_{d,e}$, with e indicating the index of the primitive in the QS ($e \in \{0, 1, ..., E\}$) and d indicating the considered derivative ($d \geq 0$). Permitted symbol values are -1, 0 +1, and ? and correspond to a positive, zero, negative sign, and undefined or unknown sign. The symbols are combined into a sign symbol matrix, $\mathbf{S}$, with row corresponds to the primitives in the qualitative sequence and columns corresponding to the index of the derivative (starting with 0 for the function or data value). The transitions between primitives are given as a vector $\boldsymbol{\theta} = \left[\begin{array}{cccc} \theta_1 & \theta_2 & ... & \theta_E \end{array} \right]$.

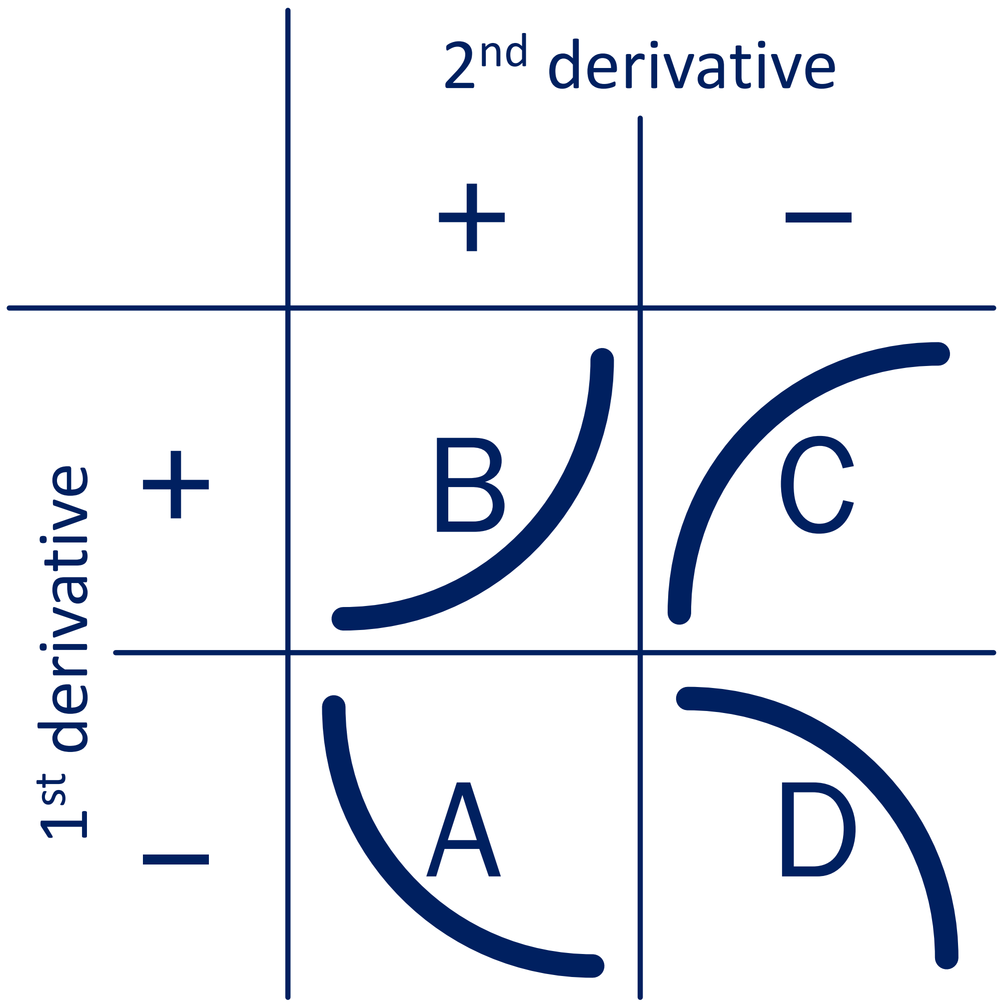

Figure 1: Primitives according to the signs of the 1st and 2nd derivative: A = anti-tonic convex, B = isotonic convex, C = isotonic concave, D = anti-tonic concave. Taken from [Villez2014].

3 Case 1: A single qualitative representation for all data series

3.1 Problem definition

A multivariate data set is given as a set of n_c univariate data series, indexed by c ($c \in \{1, 2, \ldots, C\}$). Each series is given as a vector ($\boldsymbol{y}_c$) and a corresponding vector of independent variables ($\boldsymbol{x}_c$). Repeated values within $\boldsymbol{x}_c$ correspond to repeated measurements. The sample pairs are indexed by j ($j \in \{1, 2, \ldots, J_c\}, \forall c : J_c \geq 2$). Each of these data series corresponds to measurement sets with distinct values but with the identical qualitative representation. Scalar measurements are referred to as $y_{c,j}$.The independent variable scalars are given as $x_{c,j}$. The complete set of independent (dependent) data is noted as $\overline{\mathbf{X}}$ ($\overline{\mathbf{Y}}$). All symbol definitions as used in this section are given in Table 1.

A separate shape constrained function (f_c) is estimated for each of the C data series. Importantly, each of these functions is constrained to have the same shape. In more precise terms, the shape constraints of each function are defined by the exact same qualitative representation (QR) as defined in [1]. Mathematically, this is written as the following optimization problem:

$$\min_{\boldsymbol{\eta},\boldsymbol{\theta}} g(\boldsymbol{\eta}) = g(\boldsymbol{\eta}, \overline{\mathbf{X}}, \overline{\mathbf{Y}}) \quad (1)$$

subject to:

$$\begin{aligned} \boldsymbol{\eta} &= \left[\begin{array}{cccc} \boldsymbol{\beta}_1^T & \boldsymbol{\beta}_2^T & \ldots & \boldsymbol{\beta}_C^T \end{array} \right]^T & (2) \\ \forall c \in \{1, 2, \ldots, C\} : \boldsymbol{\beta}_c &\in \Omega_c(\boldsymbol{\theta}, \mathbf{S}) & (3) \\ \boldsymbol{\theta} &\in \Theta & (4) \end{aligned}$$

with definitions and symbols as in Table 1.

Table 1: Case 1 – Definitions.

Symbol	Definition
$\boldsymbol{\beta}_c$	Parameters of the c^{th} spline function
$\boldsymbol{\epsilon}$	Optimization tolerance
$\boldsymbol{\eta}$	Parameters of the spline functions
$\boldsymbol{\theta}$	Transitions defining the enforced shape of the functions
$\boldsymbol{\theta}_{QP}$	Upper bound solution for $\boldsymbol{\theta}$
$\lambda_{c,d}$	Smoothness penalty coefficient ($\forall c \in \{1, 2, \ldots, C\}, \forall d \in \{0, 1, \ldots, D\} : \lambda_{c,d} \geq 0$)
$\Omega_c(\boldsymbol{\theta}, \mathbf{S})$	Feasible set for $\boldsymbol{\beta}_c$, which depends on $\boldsymbol{\theta}$ and $\mathbf{S}$
Θ	Feasible set for $\boldsymbol{\theta}$
Θ_k	k^{th} considered set for $\boldsymbol{\theta}$ during optimization
C	Number of functions to fit
D	Maximum considered derivative degree for shape constraints
E	Number of episodes in the qualitative representation
J_c	Number of samples in data series (c)
$\mathbf{S}$	Matrix describing the qualitative sequence, i.e. series of primitives
$\overline{\mathbf{X}}$	Independent data (argument values)
$\overline{\mathbf{Y}}$	Dependent data (measurements)
$\boldsymbol{b}$	Boundary argument values for each episode
d	Derivative index
f_c	c^{th} function
g	Objective function
h_c	c^{th} term in the objective function
j	Sample index
p_c	Power for lack-of-fit norm ($\forall c \in \{1, 2, \ldots, C\} : p_c \geq 1$)
$q_{c,d}$	Power for smoothness penalty ($\forall c \in \{1, 2, \ldots, C\}, \forall d \in \{0, 1, \ldots, D\} : q_{c,d} \geq 1$)
$s_{d,e}$	Sign for d^{th} derivative in e^{th} episode
v	Function argument
$\boldsymbol{x}_c$	Independent data vector
$\boldsymbol{y}_c$	Dependent data vector

The optimal fit of the function is defined as the sum of a convex lack-of-fit objective, formulated as a norm on the residuals, combined with a smoothness penalty, formulated by a convex sum of integrals of norms of the function value or its derivatives. As a result, the problem can be formulated in a more detailed form as follows:

$$\min_{\boldsymbol{\eta},\boldsymbol{\theta}} g(\boldsymbol{\eta}) = \sum_{c=1}^{C} \Bigg(|\boldsymbol{y}_c - f_c(\boldsymbol{\beta}_c, \boldsymbol{x}_c)|^{p_c} + \sum_{d=0}^{D} \lambda_{c,d} \cdot \int_{x_1}^{x_{J_c}} |f_c^d(\boldsymbol{\beta}_c, v)|^{q_{c,d}}\, dv \Bigg) \quad (5)$$

subject to:

$$\boldsymbol{\theta} \in \Theta \quad (6)$$

$$\boldsymbol{\eta} = \left[\begin{array}{cccc} \boldsymbol{\beta}_1^T & \boldsymbol{\beta}_2^T & \dots & \boldsymbol{\beta}_C^T \end{array}\right]^T \quad (7)$$

$$\begin{aligned} \boldsymbol{b}_c^{\mathrm{L}} &= \left[\begin{array}{cccc} b_{c,0}^{\mathrm{L}} & b_{c,1}^{\mathrm{L}} & \dots & b_{c,E}^{\mathrm{L}} \end{array}\right] \\ &= \left[\begin{array}{ccccc} x_{c,1} & \theta_1 & \theta_2 & \dots & \theta_E \end{array}\right] \quad (8) \end{aligned}$$

$$\begin{aligned} \boldsymbol{b}_c^{\mathrm{U}} &= \left[\begin{array}{cccc} b_{c,0}^{\mathrm{U}} & b_{c,1}^{\mathrm{U}} & \dots & b_{c,E}^{\mathrm{U}} \end{array}\right] \\ &= \left[\begin{array}{ccccc} \theta_1 & \theta_2 & \dots & \theta_E & x_{c,J_c} \end{array}\right] \quad (9) \end{aligned}$$

$$\forall d \in \{0, 1, \dots, D\},$$
$$\forall e \in \{0, 1, \dots, E\}:$$

$$b_{c,e}^{\mathrm{L}} \leq v \leq b_{c,e}^{\mathrm{U}} \Rightarrow \begin{cases} f_c^d(\boldsymbol{\beta}_c, v) \leq 0, & \textit{if } s_{d,e} = -1 \\ f_c^d(\boldsymbol{\beta}_c, v) = 0, & \textit{if } s_{d,e} = 0 \\ f_c^d(\boldsymbol{\beta}_c, v) \geq 0, & \textit{if } s_{d,e} = +1 \end{cases} \quad (10)$$

3.2 Solving for $\boldsymbol{\eta}$

The above problem is hard to solve to global optimality, in particular due to the apparent presence of an infinite number of shape constraint inequalities (Eq. 10). Fortunately, for specific types of functions these inequality constraints can be reformulated as a finite number of necessary and sufficient constraint equations in the function parameters. In [2] this was demonstrated for polynomial functions. In [3, 4], this is applied to univariate piece-wise polynomial functions, including spline functions. Spline functions are flexible and inherently smooth. Moreover, shape constraints can selectively be applied to segments of the spline function argument domain. Using a spline functional basis is therefore a sensible choice for shape constrained function fitting. Importantly, the above problem is then a semi-definite program (SDP) as long as values for $\boldsymbol{\theta}$ are fixed and known. Depending on the applied sign constraints and the exact objective function, the SDP is reduced to a second-order cone program (SOCP) or a quadratic program (QP) [3, 4]. For instance, shape constrained cubic spline functions can be solved at least as a SOCP.

The above problem, with given values for $\boldsymbol{\theta}$, can be split into a separate optimization of each function by solving the following problem for each value of c ($c = 1 \ldots C$):

$$\begin{aligned}\min_{\boldsymbol{\beta}_c} h_c(\boldsymbol{\beta}_c) &= |\boldsymbol{y}_{c,t} - f_c(\boldsymbol{\beta}_c, \boldsymbol{x}_{c,t})|^{p_c} \\ &\quad + \sum_{d=0}^{D} \lambda_{c,d} \cdot \int_{x_1}^{x_{J_c}} |f_c^d(\boldsymbol{\beta}_c, v)|^{q_{c,d}} \, dv \end{aligned} \tag{11}$$

subject to:

$$\begin{aligned}\boldsymbol{b}_c^{\mathrm{L}} &= \left[\begin{array}{cccc} b_{c,0}^{\mathrm{L}} & b_{c,1}^{\mathrm{L}} & \ldots & b_{c,E}^{\mathrm{L}} \end{array}\right] \\ &= \left[\begin{array}{ccccc} x_{c,1} & \theta_1 & \theta_2 & \ldots & \theta_E \end{array}\right] \end{aligned} \tag{12}$$

$$\begin{aligned}\boldsymbol{b}_c^{\mathrm{U}} &= \left[\begin{array}{cccc} b_{c,0}^{\mathrm{U}} & b_{c,1}^{\mathrm{U}} & \ldots & b_{c,E}^{\mathrm{U}} \end{array}\right] \\ &= \left[\begin{array}{ccccc} \theta_1 & \theta_2 & \ldots & \theta_E & x_{c,J_c} \end{array}\right] \end{aligned} \tag{13}$$

$$\forall d \in \{0, 1, \ldots, D\},$$
$$\forall e \in \{0, 1, \ldots, E\}:$$

$$b_{c,e}^{\mathrm{L}} \leq v \leq b_{c,e}^{\mathrm{U}} \Rightarrow \begin{cases} f_c^d(\boldsymbol{\beta}_c, v) \leq 0, & \textit{if } s_{d,e} = -1 \\ f_c^d(\boldsymbol{\beta}_c, v) = 0, & \textit{if } s_{d,e} = 0 \\ f_c^d(\boldsymbol{\beta}_c, v) \geq 0, & \textit{if } s_{d,e} = +1 \end{cases} \tag{14}$$

3.3 Solving for $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$

The full optimization problem is nonlinear in $\boldsymbol{\theta}$. However, and as shown in prior work, this kind of problem can be solved to global optimality in a deterministic manner by means of the branch-and-bound algorithm [1]. To this end, the algorithm repeatedly divides the set of feasible values for $\boldsymbol{\theta}$ into smaller subsets until convergence. For each subset, a lower and upper bound value to the objective function is computed. These bounds allow to fathom, i.e. ignore, branches in the resulting solution tree which are proved not to contain the global optimum. In what follows, the bounding procedures enabling such fathoming are explained and proven.

3.3.1 Step 1: Finding a feasible solution for $\boldsymbol{\theta}$

Any k^{th} subset, Θ_k, generated during execution of the branch-and-bound algorithm can be described as follows:

$$\forall e \in \{1, 2, \ldots, E\} \quad : \quad \theta_e^{\mathrm{L}} \leq \theta_e \leq \theta_e^{\mathrm{U}} \tag{15}$$

In addition, each feasible solution within this set satifies the following order relationship:

$$\forall e \in \{1, 2, \ldots, E-1\} \quad : \quad \theta_e \leq \theta_{e+1} \tag{16}$$

A practical way to propose a feasible solution is obtained by solving the following optimization problem subject to the above conditions (Eq. 15–16):

$$\min_{\boldsymbol{\theta}} \sum_{e=1}^{E} (\theta_e - \theta_e^{\mathrm{L}})^2 + (\theta_e - \theta_e^{\mathrm{U}})^2 \tag{17}$$

The solution, if found, is further referred to as $\boldsymbol{\theta}^{\mathrm{QP}}$. It is possible however that no solution can be found due to the fact that the set defined by Eq. 15–16 is empty. This case is dealt with separately below.

3.3.2 Step 2a: No feasible solution available

Procedure In this case, the bounding procedures are very straightforward. As in prior work, both the upper and lower bound are set to $+\infty$:

$$g^{\mathrm{L}} = g^{\mathrm{L}}(\Theta_k) = g^{\mathrm{U}} = g^{\mathrm{U}}(\Theta_k) \quad = \quad +\infty \tag{18}$$

Proof The proof of these bounds is fairly trivial. Indeed, if no feasible solution can be found $\boldsymbol{\theta}$, then there no solution can be found with any objective function value lower than $+\infty$. This automatically also defines the upper bound at the same value. This concludes the proof.

3.3.3 Step 2b: A feasible solution is found

Computing the upper and lower bounds is more involved when a feasible solution for $\boldsymbol{\theta}$, namely $\boldsymbol{\theta}^{\mathrm{QP}}$, has been found.

Upper bound

Procedure An upper bound value for the objective function is computed by replacing $\boldsymbol{\theta}$ with the proposed solution ($\boldsymbol{\theta}^{\mathrm{QP}}$) in the original problem. This means the following problem is now solved for each value of c ($c = 1 \dots C$):

$$\begin{aligned} \min_{\boldsymbol{\beta}_c} h_c(\boldsymbol{\beta}_c) &= |\boldsymbol{y}_{c,t} - f_c(\boldsymbol{\beta}_c, \boldsymbol{x}_{c,t})|^{p_c} \\ &\quad + \sum_{d=0}^{D} \lambda_{c,d} \cdot \int_{x_1}^{x_{J_c}} |f_c^d(\boldsymbol{\beta}_c, v)|^{q_{c,d}} \, dv \end{aligned} \tag{19}$$

subject to:

$$\begin{aligned} \boldsymbol{b}_c^{\mathrm{L}} &= [\ b_{c,0}^{\mathrm{L}} \quad b_{c,1}^{\mathrm{L}} \quad \dots \quad b_{c,E}^{\mathrm{L}}\ ] \\ &= [\ x_{c,1} \quad \theta_1^{\mathrm{QP}} \quad \theta_2^{\mathrm{QP}} \quad \dots \quad \theta_E^{\mathrm{QP}}\ ] \end{aligned} \tag{20}$$

$$\begin{aligned} \boldsymbol{b}_c^{\mathrm{U}} &= [\ b_{c,0}^{\mathrm{U}} \quad b_{c,1}^{\mathrm{U}} \quad \dots \quad b_{c,E}^{\mathrm{U}}\ ] \\ &= [\ \theta_1^{\mathrm{QP}} \quad \theta_2^{\mathrm{QP}} \quad \dots \quad \theta_E^{\mathrm{QP}} \quad x_{c,J_c}\ ] \end{aligned} \tag{21}$$

$$\forall d \in \{0, 1, \dots, D\},$$
$$\forall e \in \{0, 1, \dots, E\}:$$

$$b_{c,e}^{\mathrm{L}} \leq v \leq b_{c,e}^{\mathrm{U}} \ \Rightarrow \ \begin{cases} f_c^d(\boldsymbol{\beta}_c, v) \leq 0, & \textit{if } s_{d,e} = -1 \\ f_c^d(\boldsymbol{\beta}_c, v) = 0, & \textit{if } s_{d,e} = 0 \\ f_c^d(\boldsymbol{\beta}_c, v) \geq 0, & \textit{if } s_{d,e} = +1 \end{cases} \tag{22}$$

Each of the above problems is an SDP and can thus be solved to deterministic global optimality by means of interior-point algorithms. The individidual solutions for $\boldsymbol{\beta}_c$ are referred to as $\boldsymbol{\beta}_c^{\mathrm{U}}$. The vector containing all spline coefficients is referred to as $\boldsymbol{\eta}^{\mathrm{U}}$. The resulting overall objective function is an upper bound to the objective function:

$$g^{\mathrm{U}} = g^{\mathrm{U}}(\Theta_k) = \sum_{c=1}^{C} h_c(\boldsymbol{\beta}_c^{\mathrm{U}}) \tag{23}$$

Proof This optimization completes the computation of an upper bound for $g(\boldsymbol{\eta})$. The objective function value for the computed solution is indeed an upper bound since the existence of the associated solution proves that at least one solution has a value equal or lower to this value.

Lower bound

Procedure A lower bound can be computing by means of the following relaxation of the problem. For the considered subset Θ_k, one applies only those sign constraints which are applied irrespectively of which solution within the set one chooses. Practically, the entries for $\boldsymbol{\theta}$ are replaced in the original problem by their bounds as follows:

$$\begin{aligned}\min_{\boldsymbol{\beta}_c} h_c(\boldsymbol{\beta}_c) &= |\boldsymbol{y}_{c,t} - f_c(\boldsymbol{\beta}_c, \boldsymbol{x}_{c,t})|^{p_c} \\ &\quad + \sum_{d=0}^{D} \lambda_{c,d} \cdot \int_{x_1}^{x_{J_c}} |f_c^d(\boldsymbol{\beta}_c, v)|^{q_{c,d}}\ dv \end{aligned} \tag{24}$$

subject to:

$$\begin{aligned}\boldsymbol{b}_c^{\mathrm{L}} &= [\ b_{c,0}^{\mathrm{L}} \quad b_{c,1}^{\mathrm{L}} \quad \ldots \quad b_{c,E}^{\mathrm{L}}\ ] \\ &= [\ x_{c,1} \quad \theta_1^{\mathrm{U}} \quad \theta_2^{\mathrm{U}} \quad \ldots \quad \theta_E^{\mathrm{U}}\ ] \end{aligned} \tag{25}$$

$$\begin{aligned}\boldsymbol{b}_c^{\mathrm{U}} &= [\ b_{c,0}^{\mathrm{U}} \quad b_{c,1}^{\mathrm{U}} \quad \ldots \quad b_{c,E}^{\mathrm{U}}\ ] \\ &= [\ \theta_1^{\mathrm{L}} \quad \theta_2^{\mathrm{L}} \quad \ldots \quad \theta_E^{\mathrm{L}} \quad x_{c,J_c}\ ] \end{aligned} \tag{26}$$

$\forall d \in \{0, 1, \ldots, D\}$,
$\forall e \in \{0, 1, \ldots, E\}$:

$$b_{c,e}^{\mathrm{L}} \leq v \leq b_{c,e}^{\mathrm{U}} \quad \Rightarrow \quad \begin{cases} f_c^d(\boldsymbol{\beta}_c, v) \leq 0, & \textit{if } s_{d,e} = -1 \\ f_c^d(\boldsymbol{\beta}_c, v) = 0, & \textit{if } s_{d,e} = 0 \\ f_c^d(\boldsymbol{\beta}_c, v) \geq 0, & \textit{if } s_{d,e} = +1 \end{cases} \tag{27}$$

Each of the above problems is a again an SDP and can thus be solved to deterministic global optimality by means of interior-point algorithms. The individidual solutions for $\boldsymbol{\beta}_c$ are referred to as $\boldsymbol{\beta}_c^{\mathrm{L}}$. The resulting overall objective function is a lower bound to the objective function:

$$g^{\mathrm{L}} = g^{\mathrm{L}}(\Theta_k) = \sum_{c=1}^{C} h_c(\boldsymbol{\beta}_c^{\mathrm{L}}) \tag{28}$$

Proof Because the applied constraints in the modified lower bounding problem are always applied for any particular choice of $\boldsymbol{\theta}$ within the considered solution set, one can write that the feasible set for $\boldsymbol{\beta}$ in the lower bound case, $\Omega_c^{\mathrm{L}}(\Theta_k, \mathbf{S})$, includes the feasible set for any feasible proposal for $\boldsymbol{\theta}$:

$$\forall \boldsymbol{\theta} \in \Theta_k : \Omega_c(\boldsymbol{\theta}, \mathbf{S}) \subseteq \Omega_c^{\mathrm{L}}(\Theta_k, \mathbf{S}) \tag{29}$$

Given that the objective function and remaining constraints remain unchanged, it holds that the proposed procedure leads to a proven lower bound. Consider that a feasible vector $\boldsymbol{\theta}$ leads to optimized spline coefficients $\boldsymbol{\eta}_\theta$, then one can write:

$$\forall \boldsymbol{\theta} \in \Theta_k : g^{\mathrm{L}} \leq g(\boldsymbol{\eta}_\theta) \tag{30}$$

This proves the lower bound.

3.3.4 Convergence

It can be shown the lower bound solution will deliver the globally optimal solution within a considered set for the transitions in a number of special cases and only when the considered intervals define a set for $\boldsymbol{\theta}$ that does not include any spline basis knot inside its boundaries. The special cases include transitions implying a change of sign in a piece-wise linear or quadratic polynomial. In the case of cubic spline functions this corresponds to inflection points (2$^{\mathrm{nd}}$ derivative is piece-wise linear) and extrema (1$^{\mathrm{st}}$ derivative is piece-wise quadratic). This was demonstrated in [1] for the univariate case ($R = 1, C = 1$). This features makes that the upper bound solution can be improved so that it matches the lower bound solution, effectively resulting in a collapse of the bounds onto each other. As a result, the transition locations can be solved to absolute precision and to global optimality in a finite number of steps. This is limited to the single-function case ($C = 1$).

When fitting multiple functions with the same shape constraints, the above features of the optimization problem do not necessarily hold anymore. For example, the lower bound solution described above computed for a set of shape constrained cubic spline functions and for a set of transitions which are each limited to a single knot interval will ensure that implied sign changes in the first and second derivatives occur within the same knot interval for each fitted function. This does not guarantee however that the sign changes for each function occur at the exact same location within the considered intervals. As such, a collapse of the lower and upper bound is unlikely. A few convergence properties remain true however. These are given here without proof.

Conjecture 1. The lower and upper bound converge to each other with decreasing size of the considered solution sets:

$$\left(g^{\mathrm{U}}(\Theta_k) - g^{\mathrm{L}}(\Theta_k)\right)\Big|_{k\rightarrow+\infty} \rightarrow 0 \tag{31}$$

Conjecture 2. Any strictly positive tolerance ($\epsilon > 0$) for the difference between the bounds can be reached within a finite number of iterations of the branch-and-bound algorithm:

$$\exists k \in \mathbb{N}^0 : g^{\mathrm{U}}(\Theta_k) - g^{\mathrm{L}}(\Theta_k) < \epsilon \tag{32}$$

4 Case 2: Distinct qualitative representations

In a more general case, one wishes to fit functions in such a way that different QRs are applied to different functions. Solving this problem becomes intricate when the transitions for these different QRs have to satisfy equality or inequality constraints. Proofs for this general setup are under development.

References

[1] K. Villez, R. Rengaswamy, and V. Venkatasubramanian. Generalized shape constrained spline fitting for qualitative analysis of trends. *Comp. Chem. Eng.*, 58:116–134, 2013.

[2] Y. Nesterov. *High Performance Optimization, Applied Optimization*, volume 33, chapter Squared functional systems and optimization problems, pages 405–440. Kluwer Academic Publishers, 2000.

[3] D. Papp. *Optimization models for shape-constrained function estimation problems involving nonnegative polynomials and their restrictions.* PhD thesis, Rutgers University, 2011.

[4] D. Papp and F. Alizadeh. Shape-constrained estimation using nonnegative splines. *Journal of Computational and Graphical Statistics*, 23:211–231, 2014.