ABSTRACT

CHENG, HAO. Theory and algorithms for cubic L_1 splines. (Under the direction of Shu-Cherng Fang and Henry L.W. Nuttle)

In modern geometric modeling, one of the requirements for interpolants is that they "preserve shape well." Shape preservation has often been associated with preservation of monotonicity and convexity/concavity. While shape preservation cannot yet be defined quantitatively, it is generally agreed that shape preservation involves eliminating extraneous non-physical oscillation. Classical splines, which exhibit extraneous oscillation, do not "preserve shape well."

Recently, Lavery introduced a new class of cubic L_1 splines. Empirical experiment has shown that cubic L_1 splines are cable of providing C^1-smooth, shape-preserving, multi-scale interpolation of arbitrary data, including data with abrupt changes in spacing and magnitude, with no need for monotonicity or convexity constraints, node adjustment or other user input. However, the shape-preserving capability of cubic L_1 splines has not been proved theoretically. The currently available algorithm only provides an approximation to the coefficients of cubic L_1 splines.

To lay the groundwork for theoretical analysis and the development of an exact algorithm, this dissertation proposes to treat cubic L_1 spline problems in a geometric programming framework. Such a framework leads to a geometric dual problem with a linear objective function and convex quadratic constraints. It also provides a linear system for dual-to-primal conversion.

We prove that cubic L_1 splines preserve shape well, in particular, in eliminating non-physical oscillations, without review of raw data or any human intervention. We also show that cubic L_1 splines perform well for multi-scale data, as well as preserve linearity and convexity/concavity under mild conditions.

An exact algorithm based on the geometric programming model is proposed for solving cubic L_1 splines. It decomposes the geometric programming problem into several independent small-sized sub-problems and applies a specialized active set algorithm to solve the sub-problems. The algorithm is numerically stable and highly parallelizable. It requires only simple algebraic operations.

THEORY AND ALGORITHMS FOR CUBIC L_1 SPLINES

by

HAO CHENG

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

OPERATIONS RESEARCH

Raleigh

2002

APPROVED BY:

Professor Shu-Cherng Fang
Chair of Advisory Committee

Professor Henry L.W. Nuttle
Co-chair of Advisory Committee

Professor Yahya Fathi

Professor John E. Lavery

Professor Elmor L. Peterson

Professor Hien T. Tran

To my family

Biography

CHENG, HAO was born in Xuchang, China on April 14, 1971. He entered the Department of Mathematics at Peking University in 1988 and obtained his B.S. and M.S. in Mathematics in 1992 and 1995, respectively. Then he worked as a software engineer in the Pecan Information Technology Company, Peking University. Three years later, in August 1998, he joined the graduate program in Operations Research at North Carolina State University.

Acknowledgements

I would like to express my deepest gratitude and appreciation to my advisor Dr. Shu-Cherng Fang for his guidance, encouragement and support in my dissertation research. I would also like to thank the other committee members, Dr. Henry L.W. Nuttle, Dr. John E. Lavery, Dr. Elmor L. Peterson, Dr. Yahya Fathi, and Dr. Hien T. Tran for their valuable comments. My thanks also go to Dr. Xiuli Chao and Dr. Yuan-Shin Lee for their helpful suggestions.

I would like to convey my gratitude to all my fellow graduate students in FANGroup and Operations Research program, especially Shyh-Huei Chen, Shunmin Wang, Ilker Birbil, Yue Dai, Burcu Ozcam, and Yi Liao.

Finally, I would like to thank my family, especially my wife Xiao Song, for their love, understanding, encouragement and support.

This work has been kindly supported by research grant #DAAG55-98-D-0003 from the Army Research Office.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In 1946, Isaac J. Schoenberg introduced the terminology "spline" for a certain type of piecewise polynomial interpolant. Since that time, splines have been shown to be applicable and effective for a large number tasks in interpolation, approximation, numerical integration and differentiation, control theory, computer aided geometric design, fractals and numerical methods of probability theory and statistics [17]. Since the 1960s, splines have undergone rapid development. More than three hundred books, thousands of papers and more than three hundred Ph.D. dissertations related to various aspects of splines and their applications can be found in the literature [17]. Good theoretical treatments of splines can be found in [1, 19, 25], while efficient algorithms can be found in [12, 28, 29].

The most basic spline functions are the polynomial splines, now often expressed as B-splines. Exponential splines and rational splines, including non-uniform rational B-splines (NURBS), have also been developed. These "classical" splines have nice structures, excellent approximation power and are easy to manipulate, evaluate and implement on a computer. However, for arbitrary data with irregular values and irregular knot spacings, these splines have the disadvantage of often exhibiting extraneous oscillation.

In classical approximation theory, the approximation power was measured by the ability of interpolants to approximate smooth functions, typically on smoothly varying grids. In this context, polynomial splines do well in general and are, in many circumstances, provably optimal. However, over the past decade, it has become clear that the assumptions of smoothness of the underlying function and of smooth variation in the grid exclude the situations of most interest to the modern geometric modeling community, namely, minimally smooth underlying functions and arbitrary grids. In this modern context, one of the requirements for interpolants is that they "preserve shape well," a requirement that was not included in classical spline formulations. Many attempts at defining shape preservation have been made, but there is still no widely accepted definition. Shape preservation has often been associated with preservation of monotonicity and convexity/concavity. While shape preservation cannot yet be defined quantitatively, it is generally agreed that shape preservation involves eliminating extraneous oscillation. For example, in modeling natural and urban terrain, one strongly prefers that, when smooth natural terrain meets a lake with a flat surface, a flat road or a building with a vertical wall and flat top, neither the terrain, lake, road nor building are represented with permanent, nonphysical oscillation (Gibbs phenomena). Classical splines, which exhibit extraneous oscillation, do not "preserve shape well" [14].

Recent computational experiments have shown that a new class of cubic L_1 splines "provide C^1-smooth, shape-preserving, multi-scale interpolation of arbitrary data, including data with abrupt changes in spacing and magnitude, with no need for monotonicity or convexity constraints, node adjustment or other user input." [14]. While empirical observation suggests that cubic L_1 splines are very promising in geometric modeling, there has been no theory to support such an observation. It is not clear whether cubic L_1 splines behave well in general or only in special cases.

In addition to the lack of theoretical analysis, current algorithms for solving cubic L_1 splines are not considered satisfactory by the geometric modeling society. The coefficients of a cubic L_1 spline are calculated by minimizing the L_1 norm of the second derivative of the spline, which is equivalent to solving a non-differentiable convex optimization problem. Currently available non-smooth nonlinear optimization algorithms [3, 13, 27] are not efficient for calculating the coefficients of cubic L_1 splines, since they sharply increase the bandwidth of the underlying matrix and do not exploit the special structure of the minimization principle. Lavery [14, 16] proposed to calculate the coefficients of L_1 splines using a discretized version of the theoretical minimization principle. This algorithm is more efficient than others, but it is still two or more orders of magnitude more expensive than the algorithms used for classical cubic splines. Moreover, the solution obtained by the discretization approach is only an approximation of the true L_1 spline and the quality of solutions has not been studied.

To lay the groundwork for theoretical analysis and algorithm development, we propose to study cubic L_1 spline problem in a geometric programming framework. To the best of our knowledge, this is the first attempt in this research direction.

Geometric programming was originally developed for solving optimization problems in posynomial form [6]. Peterson [21, 22, 23] further generalized this approach for convex analysis. Most recent development can be found in [26]. Generalized geometric programming theory provides strong existence, uniqueness and characterization theorems, which are useful for parametric analysis and algorithm design. By formulating the univariate cubic L_1 spline problem in a generalized geometric programming setting, we are able to take advantage of the underlying structure of the problem to prove some shape-preserving properties, and eventually to develop an exact and efficient algorithm.

The rest of this dissertation is arranged as follows. In Chapter 2, the cubic L_1 spline problem is briefly described. Calculating the coefficients of cubic L_1 splines results in a non-differentiable optimization problem. This problem is re-formulated as a geometric programming problem. A geometric dual with a linear objective function and convex quadratic constraints is derived. Finally, a linear system for dual to primal conversion is established.

There are several similar problems closely related to cubic L_1 interpolating splines. One is to find cubic L_1 smoothing splines [15] by minimizing a convex combination of the l_1 norm of the residuals of the data-fitting equation and the L_1 norm of the second derivative of the splines over a finite-dimensional spline space. Numerical experiments have shown that this new class of smoothing splines perform well on preserving the shapes for arbitrary data. The proposed geometric programming approach can handle this problem too. In Chapter 3, we will show that its geometric dual has a linear objective function and a convex feasible domain, despite the fact that the objective function of the primal problem is nonlinear and non-differentiable. Then it becomes possible to perform some required sensitivity analysis.

Some shape-preserving properties of cubic L_1 splines are discussed in Chapter 4. By exploiting the geometric programming model, we see that the first derivatives of cubic L_1 splines do not depend on the data spacing. Hence, cubic L_1 splines can perform well for multi-scale data. We further prove that cubic L_1 splines preserve linearity over more than three consecutive sub-intervals. The only exception occurs at the intersection of two linear segments with different slopes. Hence, cubic L_1 splines actually eliminate any possible oscillation over linear data. In cases where cubic L_1 splines fail to preserve linearity, there does not exist any C^1-smooth interpolating function that can do so. In this sense, cubic L_1 splines are the best C^1-smooth interpolating functions for preserving linearity. For convex data, cubic L_1 splines always lie below the curve of the piecewise linear interpolation function. Therefore, they do not have any non-physical oscillation. Cubic L_1 splines also preserve convexity under certain conditions. Our conclusion is that cubic L_1 splines are extraordinarily good at eliminating extraneous non-physical oscillations.

Chapter 5 develops a continuum-based algorithm within the geometric programming framework. For a convex optimization problem like the geometric dual that we obtain, common optimization techniques [11, 30] need to solve a series of linear programming sub-problems [7]. This makes them computationally more expensive than the discretization approach [14], which requires the solution of only one linear program. We take advantage of the special structure of the geometric programming model to decompose the problem into several independent small-sized sub-problems. A specialized active set algorithm is developed to solve the sub-problems. After obtaining a dual optimal solution, one substitution process is performed to get a corresponding primal optimal solution. The overall algorithm requires only simple algebraic operations. There is no need to solve a linear program or perform any matrix operations. This algorithm has several advantages over the discretization approach. First of all, what we obtain is an exact solution. Secondly, it is numerically more stable than the interior-point algorithms. Thirdly, the algorithm is highly parallelizable for computation. By decomposing the geometric programming problem into several independent sub-problems, a small perturbation of data magnitude at a knot changes the optimal solution of only one sub-problem. Hence the data magnitude only has "local" effect, and therefore, cubic L_1 splines preserve shape very well for arbitrary data with outliers and irregular magnitudes.

Finally, in Chapter 6 we conclude this dissertation with some discussions and point out future research directions.

Chapter 2

Univariate cubic L_1 splines – a geometric programming approach

The geometric programming model of univariate cubic L_1 splines is developed in this chapter. In Section 2.1, the mathematical model for univariate cubic L_1 splines is introduced. In Section 2.2, important properties of the univariate cubic L_1 splines are investigated. In Section 2.3, the univariate cubic L_1 spline problem is formulated as a geometric programming problem. A geometric dual with a linear objective function and convex quadratic constraints is derived. A linear system for dual to primal conversion is established. Discussion and conclusion are given in Section 2.4. Finally, the details of calculating the conjugate dual are shown in Section 2.5.

2.1 Univariate cubic L_1 splines

A general setting for univariate spline interpolation can be described as follows: Let there be given a strictly monotonic partition $\Delta = \{x_i\}_0^n$ of a finite real interval $[a, b]$, such that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\,.$$

The set Δ partitions the interval $[a, b]$ into n subintervals, $I_i = [x_i, x_{i+1})$, $i = 0, 1, \ldots, n-2$, and $I_{n-1} = [x_{n-1}, x_n]$. At each knot x_i, the function value, z_i, is given. The objective is to find a piecewise polynomial function $s(x)$ to interpolate the data set $\{(x_i, z_i)\}_{i=0}^n$.

Using the notation of Schumaker [25], given a positive integer m, we define

$$\mathcal{P}_m = \left\{ p\left(x\right) \middle| p\left(x\right) = \sum_{i=1}^{m} c_i x^{i-1}, \;\; c_1, \ldots, c_m, x \in R \right\}$$

to be the space of polynomials of order m,

$$\begin{aligned} \mathcal{PP}_m\left(\Delta\right) = \quad & \{f \mid f\left(x\right) = p_i\left(x\right) \quad \text{for } x \in I_i, \\ & \text{where } p_i \in \mathcal{P}_m, i = 0, 1, \ldots, n-1\} \end{aligned}$$

to be the space of piecewise polynomials of order m with knots $x_1, \ldots, x_{n-1}$ and $C^r\left[a, b\right]$ to be the space of functions whose first rth derivatives are continuous on $[a, b]$, for a given positive integer r.

"Classical cubic splines" are calculated by minimizing the L_2 norm of the second derivative. However, computational experience has shown that classical cubic splines often have excessive "non-physical" oscillation and therefore do not preserve shape well. For this reason, other variants of cubic splines have become an active research topic. Cubic splines based on minimizing the L_p norm of the second derivative—defined as follows—have been investigated [14].

Definition 2.1 *(univariate cubic L_p splines) Let $\Delta = \{x_i\}_0^n$ be a partition of the interval $[a, b]$ and let $\{(x_i, z_i)\}_{i=0}^n$ be a given data set. A piecewise cubic polynomial $S\left(x\right)$ is called a cubic L_p spline, if, for $1 \leq p < \infty$,*

$$\begin{aligned} S\left(x\right) = \underset{z(x)}{\arg\min} \Bigg\{ \sum\nolimits_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} \left|z''\left(x\right)\right|^p dx \;\Bigg|\; & z\left(x\right) \in \mathcal{PP}_4\left(\Delta\right) \cap C^1\left[a, b\right], \\ & \textit{and } z\left(x_i\right) = z_i, i = 0, \ldots, n \Bigg\}, \end{aligned}$$

and for $p = \infty$,

$$\begin{aligned} S\left(x\right) = \underset{z(x)}{\arg\min} \Bigg\{ \max_{0 \leq i \leq n-1} \max_{x \in [x_i, x_{i+1}]} \left|z''\left(x\right)\right| \;\Bigg|\; & z\left(x\right) \in \mathcal{PP}_4\left(\Delta\right) \cap C^1\left[a, b\right], \\ & \textit{and } z\left(x_i\right) = z_i, i = 0, \ldots, n \Bigg\}. \end{aligned}$$

Classical cubic splines coincide with the cubic L_2 splines of Definition 2.1 even though the classical framework involves minimization over an infinite-dimensional function space rather than over a finite-dimensional spline space. Cubic L_p splines have continuous first-order derivatives and, for $p = 2$, also continuous second-order derivatives.

Lavery [14] introduced the class of cubic L_p splines defined above and proved the existence of the cubic L_p splines for $1 \le p \le \infty$. He also showed the uniqueness of the cubic L_p spline for $1 < p \le \infty$. The computational results presented in [14] indicate that the properties of L_p splines depend strongly on p and that, the lower p is (subject to $p \ge 1$), the better the L_p spline preserves shape. For this reason, we focus in this paper on L_1 splines.

If $S(x)$ is a univariate cubic L_1 spline, then on each subinterval $[x_i, x_{i+1}]$, there exists a cubic polynomial $S_i(x)$ such that $S(x) = S_i(x)$, $\forall x \in [x_i, x_{i+1}]$, $i = 0, \ldots, n-1$. For $i = 0, \ldots, n-1$, $S_i(x)$ can be uniquely expressed as

$$S_i(x) = p_i + q_i(x - x_i) + \frac{u_i}{2}(x - x_i)^2 + \frac{v_i}{6}(x - x_i)^3. \tag{2.1}$$

Let

$$h_i = x_{i+1} - x_i,\ i = 0, \ldots, n-1, \tag{2.2}$$

and

$$\Delta z_i = \frac{z_{i+1} - z_i}{h_i},\ i = 0, \ldots, n-1. \tag{2.3}$$

On each subinterval $[x_i, x_{i+1}]$, the second derivative of $S_i(x)$ is given by

$$S_i''(x) = u_i + v_i(x - x_i).$$

Changing the variable from x to $t = (x - x_i - \frac{h_i}{2})/h_i$, $-\frac{1}{2} \le t \le \frac{1}{2}$, we obtain

$$S_i''(t) = u_i + \frac{h_i}{2}v_i + h_i v_i t.$$

The coefficients of a univariate cubic L_1 spline are calculated by solving the optimization problem

$$\begin{aligned}
&\min_{p_i,q_i,u_i,v_i} \|S''(x)\|_1 \\
&\quad = \min_{p_i,q_i,u_i,v_i} \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} |S_i''(x)|\, dx \\
&\quad = \min_{p_i,q_i,u_i,v_i} \sum_{i=0}^{n-1} h_i \int_{-1/2}^{1/2} \left| u_i + \frac{h_i}{2}v_i + h_i v_i t \right| dt && (2.4)\\
&\quad = \min_{p_i,q_i,u_i,v_i} \sum_{i=0}^{n-1} \begin{cases} \frac{h_i}{2}\left[\left|\frac{h_i}{2}v_i\right| + \dfrac{\left(u_i + \frac{h_i}{2}v_i\right)^2}{\frac{h_i}{2}v_i} \right] & \text{if } \left|\frac{h_i}{2}v_i\right| > \left|u_i + \frac{h_i}{2}v_i\right| \\ h_i \left|u_i + \frac{h_i}{2}v_i\right| & \text{if } \left|\frac{h_i}{2}v_i\right| \le \left|u_i + \frac{h_i}{2}v_i\right| \end{cases}, && (2.5)
\end{aligned}$$

under the constraints that the spline interpolate the data and be C^1 smooth, that is,

$$p_i = z_i, \qquad i = 0, \ldots, n, \tag{2.6}$$

$$p_i + h_i q_i + \frac{h_i^2}{2} u_i + \frac{h_i^3}{6} v_i = p_{i+1}, \qquad i = 0, \ldots, n-1, \tag{2.7}$$

$$q_i + h_i u_i + \frac{h_i^2}{2} v_i = q_{i+1}, \qquad i = 0, \ldots, n-1. \tag{2.8}$$

From constraints (2.6), (2.7) and (2.8), we can solve for u_i and v_i in terms of q_i,

$$u_i = -\frac{2}{h_i}\left(2q_i + q_{i+1} - 3\Delta z_i\right), \qquad i = 0, \ldots, n-1, \tag{2.9}$$

$$v_i = \frac{6}{h_i^2}\left(q_i + q_{i+1} - 2\Delta z_i\right), \qquad i = 0, \ldots, n-1. \tag{2.10}$$

Consequently, $\|S''\|_1$ is a function of q_i. We denote it by $E(\mathbf{q})$. In this way,

$$\begin{aligned} E(\mathbf{q}) &= \|S''\|_1 \\ &= \sum_{i=0}^{n-1} E_i(q_i, q_{i+1}) \\ &= \sum_{i=0}^{n-1} \begin{cases} \frac{1}{2}\left[\left|3\left(q_i + q_{i+1} - 2\Delta z_i\right)\right| + \frac{(q_i - q_{i+1})^2}{\left|3(q_i + q_{i+1} - 2\Delta z_i)\right|}\right], & \\ & \text{if} \quad \left|3\left(q_i + q_{i+1} - 2\Delta z_i\right)\right| > \left|q_i - q_{i+1}\right|, \\ \left|q_i - q_{i+1}\right|, & \text{if} \quad \left|3\left(q_i + q_{i+1} - 2\Delta z_i\right)\right| \le \left|q_i - q_{i+1}\right|. \end{cases} \end{aligned} \tag{2.11}$$

Hence, the constrained optimization problem (2.5), (2.6), (2.7), (2.8) is equivalent to the unconstrained optimization problem

$$\min_{\mathbf{q} \in R^{n+1}} E(\mathbf{q}).$$

2.2 Properties of univariate cubic L_1 spline functional

In this section we investigate some important properties of the objective function $E(\mathbf{q})$. In particular, we prove that $E(\mathbf{q})$ is convex and almost everywhere differentiable.

The following two simple propositions from convex analysis [24] will be used to prove the convexity of $E(\mathbf{q})$.

Proposition 2.2 *If $f_i(x) : R^m \to R^1$ is a convex function and $\alpha_i \ge 0$, $\forall\, 1 \le i \le n$, then $f(x) = \sum_{i=1}^n \alpha_i f_i(x)$ is a convex function.*

Proposition 2.3 *If $f : R^m \to R^1$ is a convex function and $h : R^n \to R^m$ is an affine transformation, then the composition $f \circ h : R^n \to R^1$ is a convex function.*

Now, let us define a function $F : R^2 \to R$ by

$$F(x,y) = h\int_{-\frac{1}{2}}^{\frac{1}{2}} \left| x + \frac{h}{2}y + hyt \right| dt. \tag{2.12}$$

Note that each individual term $E_i(q_i, q_{i+1})$ of the objective function $E(\mathbf{q})$ in (2.11) can be expressed as the composition of $F(u_i, v_i)$ and an affine transformation from (u_i, v_i) to (q_i, q_{i+1}) as defined in (2.9) and (2.10). Hence, in order to prove the convexity of $E(\mathbf{q})$, it is sufficient to show that $F(x,y)$ is a convex function on R^2.

Proposition 2.4 *$F(x,y)$ is a convex function on R^2.*

Proof. *It is clear that $\left|x + \frac{h}{2}y + hty\right|$ is a convex function in terms of x and y. Hence, for any $(x_1, y_1), (x_2, y_2) \in R^2$ and any $\lambda \in [0,1]$, we have*

$$\begin{aligned}
&F(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \\
&= h\int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \lambda \left(x_1 + \frac{h}{2}y_1 + hty_1 \right) + (1-\lambda)\left(x_2 + \frac{h}{2}y_2 + hty_2 \right) \right| dt \\
&\le h\int_{-\frac{1}{2}}^{\frac{1}{2}} \lambda \left| x_1 + \frac{h}{2}y_1 + hty_1 \right| dt + h\int_{-\frac{1}{2}}^{\frac{1}{2}} (1-\lambda)\left| x_2 + \frac{h}{2}y_2 + hty_2 \right| dt \\
&= \lambda F(x_1, y_1) + (1-\lambda) F(x_2, y_2).
\end{aligned}$$

■

The following theorem is a direct consequence of Propositions 2.2, 2.3 and 2.4.

Theorem 2.5 *$E(\mathbf{q})$ is a convex function on R^{n+1}.*

For any $(x,y) \in R^2$ such that $|x| > |y|$, we notice that,

$$\frac{1}{2}|x| + \frac{1}{2}\frac{y^2}{|x|} \le \frac{1}{2}|x| + \frac{1}{2}|y| < |x|.$$

Therefore, when x goes to zero, the value of $\frac{1}{2}|x| + \frac{1}{2}\frac{y^2}{|x|}$ also goes to zero. Hence, $F(x,y)$ and the objective function $E(\mathbf{q})$ are well defined. Furthermore,

$$\lim_{|x| \to |y|} F(x,y) = |y|.$$

Hence,

Proposition 2.6 *$E_i(q_{i,}q_{i+1})$ is continuous on R^2 and $E(\mathbf{q})$ is a continuous function on R^{n+1}.*

Now we consider the continuity of the first order derivatives of $E_i(q_{i,}q_{i+1})$. Notice that the possible points of discontinuity of the first order derivatives of $E_i(q_{i,}q_{i+1})$ are given by the condition $|3(q_i + q_{i+1} - 2\Delta z_i)| = |q_i - q_{i+1}|$. By computing the gradient of $E_i(q_{i,}q_{i+1})$ on both sides of the lines defined by $|3(q_i + q_{i+1} - 2\Delta z_i)| = |q_i - q_{i+1}|$, we can prove the next result without much difficulty.

Theorem 2.7 *$E_i(q_{i,}q_{i+1}) \in C^1\left(R^2 \backslash \{\Delta z_i, \Delta z_i\}\right)$ and $E(\mathbf{q})$ is almost everywhere differentiable.*

2.3 Geometric programming approach

2.3.1 Generalized geometric programming

Geometric programming is an optimization theory with a wide range of applications. In this section, we briefly introduce the basic theory of generalized geometric programming for unconstrained optimization problems. A thorough treatment can be found in [21, 22].

Primal geometric program

In generalized geometric programming, the primal problem is to find a minimizer of a real-valued function $\mathfrak{g}(\mathbf{x})$ over a given subset $\mathfrak{F}$, which is the intersection of the function domain $\mathfrak{C} \subseteq R^n$ and a cone $\mathfrak{X} \subseteq R^n$, i.e.,

$$\text{(Primal)} \qquad \begin{cases} \min \mathfrak{g}(\mathbf{x}) \\ \mathbf{x} \in \mathfrak{C} \cap \mathfrak{X} \end{cases}. \tag{2.13}$$

Conjugate transform

Definition 2.8 *(conjugate transform) Given a function $w(\mathbf{z})$ with domain $W \subseteq R^n$, the conjugate transform of $w(\mathbf{z})$ is a function $\omega(\boldsymbol{\zeta})$ with domain $\Omega \subseteq R^n$, where*

$$\Omega = \left\{ \boldsymbol{\zeta} \in R^n \,\middle|\, \sup_{\mathbf{z} \in W} \left[\langle \boldsymbol{\zeta}, \mathbf{z} \rangle - w(\mathbf{z})\right] < +\infty \right\},$$

and

$$\omega(\boldsymbol{\zeta}) = \sup_{\mathbf{z}\in W} [\langle \boldsymbol{\zeta}, \mathbf{z}\rangle - w(\mathbf{z})], \ \ \forall \boldsymbol{\zeta} \in \Omega.$$

For a given function w, if the domain of its conjugate transform is empty, we say that its conjugate transform "does not exist". It is known that the conjugate transform of a convex function always exists.

Theorem 2.9 *[22] Given a function $w(\mathbf{z})$ with domain $W \subseteq R^n$, if $w(\mathbf{z})$ is a convex function and W is a nonempty convex set, then there exists a conjugate transform of $w(\mathbf{z})$.*

The above theorem and the definition of the conjugate transform give us the following important inequality:

Theorem 2.10 *[22] (conjugate inequality) For each $\mathbf{z} \in W$ and each $\boldsymbol{\zeta} \in \Omega$,*

$$\langle \boldsymbol{\zeta}, \mathbf{z}\rangle \leq w(\mathbf{z}) + \omega(\boldsymbol{\zeta}), \tag{2.14}$$

with equality holding if and only if $\boldsymbol{\zeta} \in \partial w(\mathbf{z})$.

Dual geometric program

Given a function $\mathfrak{g}(\mathbf{x})$ over domain $\mathfrak{C}$, denoted by $\mathfrak{g} : \mathfrak{C}$, the primal problem is given by (2.13). The conjugate transform of $\mathfrak{g} : \mathfrak{C}$ is $\mathfrak{h}(\mathbf{y})$ with domain $\mathfrak{D}$, denoted by $\mathfrak{h} : \mathfrak{D}$, where

$$\mathfrak{D} = \left\{ \mathbf{y} \in R^n \,\middle|\, \sup_{\mathbf{x}\in\mathfrak{C}} [\langle \mathbf{y}, \mathbf{x}\rangle - \mathfrak{g}(\mathbf{x})] < +\infty \right\}, \tag{2.15}$$

and

$$\mathfrak{h}(\mathbf{y}) = \sup_{\mathbf{x}\in\mathfrak{C}} [\langle \mathbf{y}, \mathbf{x}\rangle - \mathfrak{g}(\mathbf{x})], \ \ \forall \mathbf{y} \in \mathfrak{D}. \tag{2.16}$$

The feasible region of the primal problem is the intersection of domain $\mathfrak{C}$ with some cone $\mathfrak{X} \subseteq R^n$. Let $\mathfrak{Y}$ be the dual cone of $\mathfrak{X}$, which is defined by

$$\mathfrak{Y} = \{\mathbf{y} \in R^n \mid \langle \mathbf{x}, \mathbf{y}\rangle \geq 0, \ \ \forall \mathbf{x} \in \mathfrak{X}\}. \tag{2.17}$$

Then the dual problem becomes

$$\text{(Dual)} \qquad \begin{cases} \min \mathfrak{h}(\mathbf{y}) \\ \mathbf{y} \in \mathfrak{D} \cap \mathfrak{Y} \end{cases}. \tag{2.18}$$

Optimality conditions

Theorem 2.11 *[22] (optimality conditions)* $\mathbf{x}^*$ *and* $\mathbf{y}^*$ *are optimal solutions to the primal problem (2.13) and the dual problem (2.18), respectively, if and only if*

$$\begin{array}{ll} \text{(I)} & \mathbf{x}^* \in \mathfrak{C} \cap \mathfrak{X},\ \mathbf{y}^* \in \mathfrak{D} \cap \mathfrak{Y}; \\ \text{(II)} & \langle \mathbf{x}^*, \mathbf{y}^* \rangle = 0; \\ \text{(III)} & \mathbf{y}^* \in \partial \mathfrak{g}\left(\mathbf{x}^*\right). \end{array}$$

Optimality condition (I) indicates primal and dual feasibility. Optimality condition (II) is called the "orthogonality condition". If the primal cone $\mathfrak{X}$ is actually a vector space, then its dual cone $\mathfrak{Y} = \mathfrak{X}^{\perp}$. Hence, the orthogonality condition is automatically satisfied and can be omitted. Optimality condition (III) is called the "subgradient condition". When both function $\mathfrak{g} : \mathfrak{C}$ and cone $\mathfrak{X}$ are convex and closed, it is known that the primal problem (2.13) becomes the dual of the dual problem (2.18). In this case, the duality between primal problem (2.13) and the dual problem (2.18) is symmetric; so the optimality condition (III) is equivalent to

$$\text{(III}a\text{)} \quad \mathbf{x}^* \in \partial \mathfrak{h}\left(\mathbf{y}^*\right) \text{ and } \mathbf{y}^* \in \partial \mathfrak{g}\left(\mathbf{x}^*\right).$$

Theorem 2.12 *[22] If* $\mathbf{x}$ *and* $\mathbf{y}$ *are feasible solutions to the primal problem (2.13) and the dual problem (2.18), respectively, then*

$$0 \leq \mathfrak{g}\left(\mathbf{x}\right) + \mathfrak{h}\left(\mathbf{y}\right),$$

with equality holding if and only if the optimality conditions (II) and (III) are satisfied. In this case, $\mathbf{x}$ *and* $\mathbf{y}$ *are optimal solutions to the primal problem (2.13) and the dual problem (2.18), respectively.*

Theorem 2.13 *[21] If the dual problem (2.18) has a feasible solution* $\mathbf{y}^* \in ri\left(\mathfrak{D}\right) \bigcap ri\left(\mathbf{y}\right)$*, and if* $\inf_{\mathbf{y} \in \mathfrak{D} \cap \mathfrak{Y}} \mathfrak{h}\left(\mathbf{y}\right) > -\infty$*, then the primal problem (2.13) has a nonempty solution set and*

$$0 = \inf_{\mathbf{x} \in \mathfrak{C} \cap \mathfrak{X}} \mathfrak{g}\left(\mathbf{x}\right) + \inf_{\mathbf{y} \in \mathfrak{D} \cap \mathfrak{Y}} \mathfrak{h}\left(\mathbf{y}\right).$$

2.3.2 Geometric programming approach for univariate cubic L_1 splines

Primal problem

Denote

$$\mathbf{x}^i = (p_i, q_i, u_i, v_i), \quad i = 0, \ldots, n-1,$$

$$\mathbf{x}^n = (p_n, q_n),$$

and

$$\mathbf{x} = \left(\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{n-1}, \mathbf{x}^n\right)$$

$$= (p_0, q_0, u_0, v_0; p_1, q_1, u_1, v_{1;}, \ldots; p_{n-1}, q_{n-1}, u_{n-1,} v_{n-1}; p_n, q_n).$$

For $i = 0, \ldots, n-1$, $\mathbf{x}^i$ is defined on a subset

$$\mathfrak{C}_i = \{z_i\} \times R^1 \times R^1 \times R^1 \subset R^4.$$

For $i = n$, $\mathbf{x}^n$ is defined on a subset

$$\mathfrak{C}_n = \{z_n\} \times R^1 \subset R^2.$$

Therefore, $\mathbf{x}$ is defined on the Cartesian product of all $\mathfrak{C}_i$, $i = 0, \ldots, n$, i.e.,

$$\mathfrak{C} = \prod_{i=0}^{n} \mathfrak{C}_i \subset R^{4n+2}.$$

In this case, the first set of constraints, (2.6), are automatically satisfied for any $\mathbf{x} \in \mathfrak{C}$. Moreover, constraints (2.7) and (2.8) can be treated as a cone constraint of the primal problem, since they can be expressed as

$$\mathcal{A}\mathbf{x} = 0,$$

where $\mathcal{A}$ is the coefficient matrix of (2.7) and (2.8), i.e.,

$$
\mathcal{A} = \begin{bmatrix}
1 & h_0 & \frac{h_0^2}{2} & \frac{h_0^3}{6} & -1 & & & & & & & & & & & & \\
& 1 & h_0 & \frac{h_0^2}{2} & 0 & -1 & & & & & & & & & & & \\
& & & & 1 & h_1 & \frac{h_1^2}{2} & \frac{h_1^3}{6} & -1 & & & & & & & & \\
& & & & & 1 & h_1 & \frac{h_1^2}{2} & 0 & -1 & & & & & & & \\
& & & & & & & & & \cdots & & & & & & & \\
& & & & & & & & & & \cdots & 1 & h_{n-1} & \frac{h_{n-1}^2}{2} & \frac{h_{n-1}^3}{6} & -1 & \\
& & & & & & & & & & \cdots & & 1 & h_{n-1} & \frac{h_{n-1}^2}{2} & 0 & -1
\end{bmatrix}_{2n\times(4n+2)} \tag{2.19}
$$

We define the cone $\mathfrak{X}$ to be the null space of matrix $\mathcal{A}$ for the primal problem.

The primal objective function now becomes

$$
\mathfrak{g}(\mathbf{x}) = \sum_{i=0}^{n-1} \mathfrak{g}_i\left(\mathbf{x}^i\right), \quad \mathbf{x} \in \mathfrak{C} \cap \mathfrak{X},
$$

where

$$
\begin{aligned}
&\mathfrak{g}_i\left(\mathbf{x}^i\right) \\
&= \mathfrak{g}_i\left(p_i, q_i, u_i, v_i\right) \\
&= \begin{cases}
\frac{h_i}{2}\left[\left|\frac{h_i}{2}v_i\right| + \frac{u_i+\frac{h_i}{2}v_i}{\frac{h_i}{2}v_i}^{2}\right], & \text{if} \quad \left|\frac{h_i}{2}v_i\right| > \left|u_i + \frac{h_i}{2}v_i\right|, \\
h_i\left|u_i + \frac{h_i}{2}v_i\right|, & \text{if} \quad \left|\frac{h_i}{2}v_i\right| \le \left|u_i + \frac{h_i}{2}v_i\right|.
\end{cases}
\end{aligned}
$$

Consequently, the primal problem becomes

$$
\text{(Primal)} \qquad \begin{cases} \min \mathfrak{g}(\mathbf{x}) \\ \mathbf{x} \in \mathfrak{C} \cap \mathfrak{X} \end{cases}. \tag{2.20}
$$

For this primal problem, there are $4n+2$ unknown variables and $(n+1)+n+n = 3n+1$ constraints. Therefore, there are $n+1$ degrees of freedom.

Recall from (2.9) and (2.10) that, for any $\mathbf{x} \in \mathfrak{X}$, u_i and v_i can be expressed in terms of the q_i:

$$
u_i = -\frac{2}{h_i}\left(2q_i + q_{i+1} - 3\Delta z_i\right), \qquad i = 0, \ldots, n-1,
$$

$$
v_i = \frac{6}{h_i^2}\left(q_i + q_{i+1} - 2\Delta z_i\right), \qquad i = 0, \ldots, n-1.
$$

Hence, any feasible solution $\mathbf{x} = \left(\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{n-1}, \mathbf{x}^n\right)$ of the primal problem can be expressed in terms of the $n+1$ variables q_i, $i = 0, \ldots, n$:

$$\begin{aligned} \mathbf{x}^i &= (p_i, q_i, u_i, v_i) \\ &= \left(z_i, q_i, -\frac{2}{h_i}\left(2q_i + q_{i+1} - 3\Delta z_i\right), \frac{6}{h_i^2}\left(q_i + q_{i+1} - 2\Delta z_i\right)\right), \qquad i = 0, \ldots, n-1, \end{aligned} \tag{2.21}$$

$$\mathbf{x}^n = (p_n, q_n) = (z_n, q_n)\,. \tag{2.22}$$

Dual problem

Denote

$$\mathbf{y}^i = \left(y_1^i, y_2^i, y_3^i, y_4^i\right), \qquad i = 0, \ldots, n-1,$$

$$\mathbf{y}^n = \left(y_1^n, y_2^n\right),$$

and

$$\begin{aligned} \mathbf{y} &= \left(\mathbf{y}^0, \mathbf{y}^1, \ldots, \mathbf{y}^{n-1}, \mathbf{y}^n\right) \\ &= \left(y_1^0, y_2^0, y_3^0, y_4^0; y_1^1, y_2^1, y_3^1, y_4^1; \ldots, y_1^{n-1}, y_2^{n-1}, y_3^{n-1}, y_4^{n-1}; y_1^n, y_2^n\right). \end{aligned}$$

The conjugate transform of $\mathfrak{g}(\mathbf{x}) : \mathfrak{C}$ is denoted by $\mathfrak{h}(\mathbf{y}) : \mathfrak{D}$. From Section 2.5, we can derive

$$\mathfrak{h}(\mathbf{y}) = \sum_{i=0}^{n} \mathfrak{h}_i\left(\mathbf{y}^i\right) = \sum_{i=0}^{n} z_i y_1^i,$$

and its domain

$$\mathfrak{D} = \prod_{i=0}^{n} \mathfrak{D}_i \subset R^{4n+2},$$

where, for $i = 0, \ldots, n-1$,

$$\mathfrak{D}_i = R \times \{0\} \times \Theta_i,$$

with

$$\begin{aligned} \Theta_i = \quad & \left\{\left(y_3^i, y_4^i\right) \in R^2 \middle| -h_i \le y_3^i \le h_i, \right. \\ & \left. -\tfrac{h_i^2}{2} + \tfrac{1}{4}\left(y_3^i + h_i\right)^2 \le y_4^i \le \tfrac{h_i^2}{2} - \tfrac{1}{4}\left(y_3^i - h_i\right)^2\right\}, \end{aligned} \tag{2.23}$$

and for $i = n$,

$$\mathfrak{D}_n = R \times \{0\}\,.$$

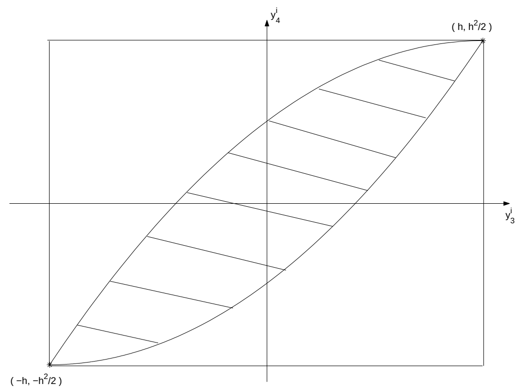

Figure 2.1: The graph of Θ_i

Therefore, the dual problem becomes

$$\begin{cases} \min \mathfrak{h}\left(\mathbf{y}\right) \\ \mathbf{y} \in \mathfrak{D} \cap \mathfrak{Y} \end{cases}, \tag{2.24}$$

where the dual cone $\mathfrak{Y}$ is the row space of the matrix

$$\mathcal{A} = \begin{bmatrix} 1 & h_0 & \frac{h_0^2}{2} & \frac{h_0^3}{6} & -1 & & & & & & & & & & & \\ & 1 & h_0 & \frac{h_0^2}{2} & 0 & -1 & & & & & & & & & & \\ & & & & 1 & h_1 & \frac{h_1^2}{2} & \frac{h_1^3}{6} & -1 & & & & & & & \\ & & & & & 1 & h_1 & \frac{h_1^2}{2} & 0 & -1 & & & & & & \\ & & & & & & & & & \cdots & & & & & & \\ & & & & & & & & & & \cdots & 1 & h_{n-1} & \frac{h_{n-1}^2}{2} & \frac{h_{n-1}^3}{6} & -1 & \\ & & & & & & & & & & \cdots & & 1 & h_{n-1} & \frac{h_{n-1}^2}{2} & 0 & -1 \end{bmatrix}_{2n\times(4n+2)}$$

Let the row vectors of $\mathcal{A}$ be $\boldsymbol{\xi}_i$ and $\boldsymbol{\eta}_i$, $i = 0, \ldots, n-1$, i.e.,

$$\mathcal{A} = \begin{bmatrix} \boldsymbol{\xi}_0 \\ \boldsymbol{\eta}_0 \\ \boldsymbol{\xi}_1 \\ \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\xi}_{n-1} \\ \boldsymbol{\eta}_{n-1} \end{bmatrix}_{2n\times(4n+2)}$$

For each $\mathbf{y} \in \mathfrak{Y}$, since $\mathfrak{Y}$ is the row space of $\mathcal{A}$, $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \sum_{i=0}^{n-1} \left(\alpha_i \boldsymbol{\xi}_i + \beta_i \boldsymbol{\eta}_i\right), \ \alpha_i, \beta_i \in R,$$

which can be explicitly calculated by

$$\begin{aligned} y_1^0 &= \alpha_0, \\ y_1^i &= -\alpha_{i-1} + \alpha_i, \ \ i = 1, \ldots, n-1, \\ y_1^n &= -\alpha_{n-1}, \end{aligned} \tag{2.25}$$

$$\begin{aligned} y_2^0 &= h_0\alpha_0 + \beta_0, \\ y_2^i &= -\beta_{i-1} + h_i\alpha_i + \beta_i, \ i = 1, \ldots, n-1, \\ y_2^n &= -\beta_{n-1}, \end{aligned} \tag{2.26}$$

$$y_3^i = \frac{h_i^2}{2}\alpha_i + h_i\beta_i, \ \ i = 0, \ldots, n-1, \tag{2.27}$$

$$y_4^i = \frac{h_i^3}{6}\alpha_i + \frac{h_i^2}{2}\beta_i, \ \ i = 0, \ldots, n-1. \tag{2.28}$$

Hence,

$$\begin{aligned} \mathfrak{h}(\mathbf{y}) &= \sum_{i=0}^{n} z_i y_1^i \\ &= z_0\alpha_0 + \sum_{i=1}^{n-1} z_i\left(-\alpha_{i-1} + \alpha_i\right) + z_n\left(-\alpha_{n-1}\right) \\ &= \sum_{i=0}^{n-1} \left(z_i - z_{i+1}\right)\alpha_i. \end{aligned}$$

The dual problem actually becomes the following optimization problem with a linear objective function over some linear and convex quadratic constraints:

$$
\begin{aligned}
\min_{\alpha_i,\beta_i} \quad & \textstyle\sum_{i=0}^{n-1} (z_i - z_{i+1})\,\alpha_i \\
\text{s.t.} \quad & h_0\alpha_0 + \beta_0 = 0, \\
& -\beta_{i-1} + h_i\alpha_i + \beta_i = 0, \quad i = 1,\dots,n-1, \\
& -\beta_{n-1} = 0, \\
& -h_i \le \alpha_i \tfrac{h_i^2}{2} + \beta_i h_i \le h_i, \quad i = 0,\dots,n-1, \\
& -\tfrac{h_i^2}{2} + \tfrac{1}{4}\left(\alpha_i \tfrac{h_i^2}{2} + \beta_i h_i + h_i\right)^2 \le \alpha_i \tfrac{h_i^3}{6} + \beta_i \tfrac{h_i^2}{2} \le \tfrac{h_i^2}{2} - \tfrac{1}{4}\left(\alpha_i \tfrac{h_i^2}{2} + \beta_i h_i - h_i\right)^2, \quad i = 0,\dots,n-1,
\end{aligned}
\tag{2.29}
$$

where α_i and β_i, $i = 0,\dots,n-1$, are unknown variables. In this setting, we have $2n$ unknown variables, $n+1$ linear equation constraints, $2n$ linear inequality constraints, and $2n$ quadratic inequality constraints.

Note that the set of constraints $-h_i \le \alpha_i \frac{h_i^2}{2} + \beta_i h_i \le h_i$ is redundant. Moreover, the last set of constraints can be simplified in terms of $\mathbf{y}$, i.e.,

$$
\left|4y_4^i - 2h_i y_3^i\right| \le (h_i)^2 - \left(y_3^i\right)^2, \quad i = 0,\dots,n-1.
$$

Using (2.27) and (2.28), we have

$$
\left|\frac{h_i}{3}\alpha_i\right| \le 1 - \left(\frac{h_i}{2}\alpha_i + \beta_i\right)^2, \quad i = 0,\dots,n-1.
$$

From the equality constraints, we can express vector $\boldsymbol{\alpha}$ as a linear function of vector $\boldsymbol{\beta}$, i.e.,

$$
\begin{aligned}
\alpha_0 &= -\tfrac{\beta_0}{h_0}, \\
\alpha_i &= \tfrac{\beta_{i-1} - \beta_i}{h_i}, \ i = 1,\dots,n-1.
\end{aligned}
\tag{2.30}
$$

This further simplifies the last set of constraints to be

$$
|\beta_0| \le \frac{\sqrt{40} - 2}{3} \approx 1.44151844,
$$

$$
\frac{1}{3}\left|\beta_{i-1} - \beta_i\right| \le 1 - \frac{1}{4}\left(\beta_{i-1} + \beta_i\right)^2, \quad i = 1,\dots,n-1,
$$

$$
\beta_{n-1} = 0.
$$

Thus, the dual problem (2.29) can be simplified as

$$\begin{cases} \min\limits_{\beta_i} & \sum\limits_{i=0}^{n-2} (\Delta z_i - \Delta z_{i+1})\, \beta_i + \Delta z_{n-1} \beta_{n-1} = \sum\limits_{i=0}^{n-2} (\Delta z_i - \Delta z_{i+1})\, \beta_i \\ \text{s.t.} & |\beta_0| \le \frac{\sqrt{40}-2}{3}, \\ & \frac{1}{3} |\beta_{i-1} - \beta_i| \le 1 - \frac{1}{4} (\beta_{i-1} + \beta_i)^2, \quad i = 1, \dots, n-1, \\ & \beta_{n-1} = 0. \end{cases} \tag{2.31}$$

In this model, the last variable β_{n-1} is fixed, hence, can be dropped. From $\beta_{n-1} = 0$ and the last quadratic constraint $\frac{1}{3} |\beta_{n-2} - \beta_{n-1}| \le 1 - \frac{1}{4} (\beta_{n-2} + \beta_{n-1})^2$, it is not difficult to see that $|\beta_{n-2}| \le \frac{\sqrt{40}-2}{3}$. Therefore, by defining

$$f(x) = |x| - \frac{\sqrt{40}-2}{3}, \tag{2.32}$$

and

$$g(x, y) = \frac{1}{3} |x - y| - 1 + \frac{1}{4} (x + y)^2, \tag{2.33}$$

we can further reduce the dual problem by one dimension and express it in a symmetric way

$$\text{(D)} \quad \begin{cases} \min\limits_{\beta_i} & \sum\limits_{i=0}^{n-2} b_i \beta_i = \langle \mathbf{B}, \boldsymbol{\beta} \rangle \\ \text{s.t.} & f(\beta_0) \le 0, \\ & g(\beta_{i-1}, \beta_i) \le 0, \quad i = 1, \dots, n-2, \\ & f(\beta_{n-2}) \le 0. \end{cases}$$

where $b_i = \Delta z_i - \Delta z_{i+1}$, $i = 0, \dots, n-2$, and $\mathbf{B} = (b_0, \dots, b_{n-2})^T$. This optimization problem has $n-1$ variables β_i, $i = 0, \dots, n-2$, $2n$ constraints, and a linear objective function. Most of the constraints are convex quadratic constraints. Surprisingly, none of the given data appears in both the constraints and the objective function.

2.3.3 Dual to primal transformation

If one of the dual optimal solutions $\mathbf{y}^*$ is obtained by solving problem (D), a corresponding primal optimal solution can be obtained by the geometric programming optimality conditions

$$\begin{array}{ll} \text{(I)} & \mathbf{x}^* \in \mathfrak{C} \cap \mathfrak{X},\ \mathbf{y}^* \in \mathfrak{D} \cap \mathfrak{Y}; \\ \text{(II)} & \langle \mathbf{x}^*, \mathbf{y}^* \rangle = 0; \\ \text{(III)} & \mathbf{x}^* \in \partial \mathfrak{h}\left(\mathbf{y}^*\right). \end{array}$$

This means that the primal optimal solution $\mathbf{x}^*$ is a vector such that

$$\mathbf{x}^* \in \mathfrak{C} \cap \mathfrak{X} \cap \partial \mathfrak{h}\left(\mathbf{y}^*\right). \tag{2.34}$$

For any $\bar{\mathbf{y}} \in \mathfrak{D}$ and for any $\boldsymbol{\gamma} \in \partial \mathfrak{h}\left(\bar{\mathbf{y}}\right)$, let

$$\boldsymbol{\gamma} = \left(\gamma_1^0, \gamma_2^0, \gamma_3^0, \gamma_4^0; \gamma_1^1, \gamma_2^1, \gamma_3^1, \gamma_4^1; \ldots; \gamma_1^{n-1}, \gamma_2^{n-1}, \gamma_3^{n-1}, \gamma_4^{n-1}; \gamma_1^n, \gamma_2^n\right).$$

From the definition of subgradient, $\boldsymbol{\gamma}$ satisfies that

$$\langle \boldsymbol{\gamma}, \mathbf{y} - \bar{\mathbf{y}} \rangle \leq \mathfrak{h}\left(\mathbf{y}\right) - \mathfrak{h}\left(\bar{\mathbf{y}}\right),\ \forall \mathbf{y} \in \mathfrak{D}. \tag{2.35}$$

Since y_2^i can only be zero, γ_2^i can be any real number. Assume $\mathbf{y}$ is a vector whose elements are the same as the given vector $\bar{\mathbf{y}}$, except for the component $y_1^i = \bar{y}_1^i + \delta$, where δ can be any real number. Then

$$\langle \boldsymbol{\gamma}, \mathbf{y} - \bar{\mathbf{y}} \rangle = \gamma_1^i \delta,$$

$$\mathfrak{h}\left(\mathbf{y}\right) - \mathfrak{h}\left(\bar{\mathbf{y}}\right) = z_i \delta.$$

Since δ can be either a positive or a negative real number, we have $\gamma_1^i = z_i$.

Assume $\mathbf{y}$ is a vector whose elements are the same as the given vector $\bar{\mathbf{y}}$, except for the components y_3^i and y_4^i. Then

$$\langle \boldsymbol{\gamma}, \mathbf{y} - \bar{\mathbf{y}} \rangle = \gamma_3^i \left(y_3^i - \bar{y}_3^i\right) + \gamma_4^i \left(y_4^i - \bar{y}_4^i\right),$$

$$\mathfrak{h}\left(\mathbf{y}\right) - \mathfrak{h}\left(\bar{\mathbf{y}}\right) = 0.$$

Formula (2.35) now becomes

$$\gamma_3^i \left(y_3^i - \bar{y}_3^i\right) + \gamma_4^i \left(y_4^i - \bar{y}_4^i\right) \leq 0,\ \ \forall \left(y_3^i, y_4^i\right) \in \Theta_i. \tag{2.36}$$

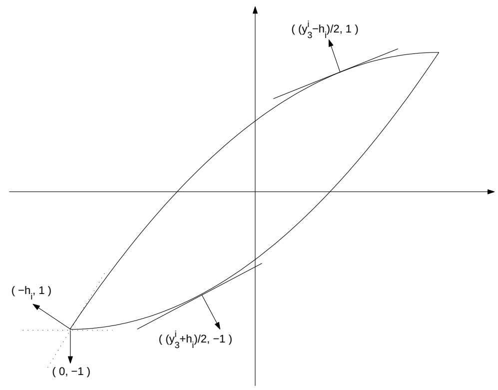

Figure 2.2: The normal cone of Θ_i at the boundary points

If we consider the above condition in the two-dimensional $\left(y_3^i, y_4^i\right)$ space, then $\left(\gamma_3^i, \gamma_4^i\right)$ must lie in the normal cone of Θ_i at point $\left(\bar{y}_3^i, \bar{y}_4^i\right)$.

Recall that $\left(y_3^i, y_4^i\right)$ is defined on the set

$$\begin{aligned}\Theta_i = \quad & \Big\{\left(y_3^i, y_4^i\right) \in R^2 \Big| -h_i \leq y_3^i \leq h_i, \\ & -\tfrac{h_i^2}{2} + \tfrac{1}{4}\left(y_3^i + h_i\right)^2 \leq y_4^i \leq \tfrac{h_i^2}{2} - \tfrac{1}{4}\left(y_3^i - h_i\right)^2\Big\},\end{aligned} \tag{2.37}$$

which is a convex set bounded by two quadratic curves:

$$C_1^i: \quad y_4^i = -\frac{h_i^2}{2} + \frac{1}{4}\left(y_3^i + h_i\right)^2,$$

and

$$C_2^i: \quad y_4^i = \frac{h_i^2}{2} - \frac{1}{4}\left(y_3^i - h_i\right)^2.$$

If $\left(\bar{y}_3^i, \bar{y}_4^i\right)$ is an interior point of the set Θ_i, then obviously

$$\left(\gamma_3^i, \gamma_4^i\right) = (0, 0).$$

If $\left(\bar{y}_3^i, \bar{y}_4^i\right)$ is a boundary point on the curve C_1^i but not on the curve C_2^i, then one normal vector of Θ_i is $\left(\frac{\bar{y}_3^i + h_i}{2}, -1\right)$. Hence

$$\begin{pmatrix} \gamma_3^i \\ \\ \gamma_4^i \end{pmatrix} = \lambda_i \begin{pmatrix} \frac{\bar{y}_3^i + h_i}{2} \\ \\ -1 \end{pmatrix}, \text{ for some } \lambda_i \geq 0.$$

If $\left(\bar{y}_3^i, \bar{y}_4^i\right)$ is a boundary point on the curve C_2^i but not on the curve C_1^i, then one normal vector of Θ_i

is $\left(\frac{\bar{y}_3^i - h_i}{2}, 1\right)$ and

$$\begin{pmatrix} \gamma_3^i \\ \\ \gamma_4^i \end{pmatrix} = \mu_i \begin{pmatrix} \frac{\bar{y}_3^i - h_i}{2} \\ \\ 1 \end{pmatrix}, \text{ for some } \mu_i \geq 0.$$

If $\left(\bar{y}_3^i, \bar{y}_4^i\right)$ is a boundary point on both the curve C_1^i and the curve C_2^i, then the normal cone is the cone spanned by vectors $\left(\frac{\bar{y}_3^i + h_i}{2}, -1\right)$ and $\left(\frac{\bar{y}_3^i - h_i}{2}, 1\right)$, i.e.,

$$\begin{pmatrix} \gamma_3^i \\ \\ \gamma_4^i \end{pmatrix} = \lambda_i \begin{pmatrix} \frac{\bar{y}_3^i + h_i}{2} \\ \\ -1 \end{pmatrix} + \mu_i \begin{pmatrix} \frac{\bar{y}_3^i - h_i}{2} \\ \\ 1 \end{pmatrix}, \text{ for some } \lambda_i, \mu_i \geq 0.$$

In summary, γ_3^i and γ_4^i should be in the form of

$$\begin{pmatrix} \gamma_3^i \\ \\ \gamma_4^i \end{pmatrix} = \lambda_i \begin{pmatrix} c_{31}^i \\ \\ c_{41}^i \end{pmatrix} + \mu_i \begin{pmatrix} c_{32}^i \\ \\ c_{42}^i \end{pmatrix}, \text{ for some } \lambda_i, \mu_i \geq 0,$$

where $\left(c_{31}^i, c_{41}^i\right) = \left(\frac{y_3^i + h_i}{2}, -1\right)$, if $\left(\bar{y}_3^i, \bar{y}_4^i\right)$ is on the curve C_1^i, otherwise $\left(c_{31}^i, c_{41}^i\right) = (0, 0)$; and $\left(c_{32}^i, c_{42}^i\right) = \left(\frac{y_3^i - h_i}{2}, 1\right)$, if $\left(\bar{y}_3^i, \bar{y}_4^i\right)$ is on the curve C_2^i, otherwise $\left(c_{32}^i, c_{42}^i\right) = (0, 0)$.

Therefore, we know that any $\boldsymbol{\gamma} \in \partial \mathfrak{h}(\bar{\mathbf{y}})$ must be of the form

$$\begin{cases} \gamma_1^i = z_i, \\ \gamma_2^i \text{ unrestricted}, \\ \gamma_3^i = \lambda_i c_{31}^i + \mu_i c_{32}^i, \\ \gamma_4^i = \lambda_i c_{41}^i + \mu_i c_{42}^i, \end{cases} \quad \text{for some } \lambda_i, \mu_i \geq 0. \tag{2.38}$$

On the other hand, it is easy to show that any vector $\boldsymbol{\gamma}$ of the above form is a subgradient of $\mathfrak{h}(\bar{\mathbf{y}})$, since

$$\begin{aligned}
&\langle \boldsymbol{\gamma}, \mathbf{y} - \bar{\mathbf{y}} \rangle \\
&= \sum_{i=0}^{n} z_i \left(y_1^i - \bar{y}_1^i\right) + \sum_{i=0}^{n-1} \left[\gamma_3^i \left(y_3^i - \bar{y}_3^i\right) + \gamma_4^i \left(y_4^i - \bar{y}_4^i\right)\right] \\
&= \left[\mathfrak{h}(\mathbf{y}) - \mathfrak{h}(\bar{\mathbf{y}})\right] + \sum_{i=0}^{n-1} \left[\gamma_3^i \left(y_3^i - \bar{y}_3^i\right) + \gamma_4^i \left(y_4^i - \bar{y}_4^i\right)\right] \\
&\leq \mathfrak{h}(\mathbf{y}) - \mathfrak{h}(\bar{\mathbf{y}}).
\end{aligned}$$

Thus, we have an explicit expression for $\partial \mathfrak{h}(\bar{\mathbf{y}})$ at a given vector $\bar{\mathbf{y}}$.

Recall from the geometric programming primal problem that any primal feasible solution

$\mathbf{x} \in \mathfrak{C} \cap \mathfrak{X}$ can be expressed in terms of variables q_i, $i = 0, \ldots, n$, i.e.,

$$\mathbf{x} = \left(\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{n-1}, \mathbf{x}^n\right),$$

where

$$\begin{aligned}
\mathbf{x}^i &= (p_i, q_i, u_i, v_i) \\
&= \left(z_i, q_i, -\frac{2}{h_i}\left(2q_i + q_{i+1} - 3\Delta z_i\right), \frac{6}{h_i^2}\left(q_i + q_{i+1} - 2\Delta z_i\right)\right), \qquad i = 0, \ldots, n-1,
\end{aligned}$$

$$\mathbf{x}^n = (p_n, q_n) = (z_n, q_n).$$

Putting this expression together with (2.38), any vector $\mathbf{x}$ is in the set $\mathfrak{C} \cap \mathfrak{X} \cap \partial\mathfrak{h}(\mathbf{y})$ if and only if

$$\begin{aligned}
&-\tfrac{2}{h_i}\left(2q_i + q_{i+1} - 3\Delta z_i\right) = \lambda_i c_{31}^i + \mu_i c_{32}^i, \\
&\tfrac{6}{h_i^2}\left(q_i + q_{i+1} - 2\Delta z_i\right) = \lambda_i c_{41}^i + \mu_i c_{42}^i, \\
&\lambda_i, \mu_i \geq 0, \\
&q_i \text{ unrestricted,}
\end{aligned} \tag{2.39}$$

with q_i, $i = 0, \ldots, n$, λ_i, μ_i, $i = 0, \ldots, n-1$, being unknown variables. After obtaining a dual optimal solution $\mathbf{y}^*$, a corresponding primal optimal solution can be obtained by solving (2.39).

However, this approach has several disadvantages. First of all, there are additional operations performed to get the value of $\mathbf{y}^*$. Secondly, $\boldsymbol{\beta}^*$ is a n-dimensional vector, while $\mathbf{y}^*$ is a $(4n+2)$-dimensional vector. Hence, this approach merely increases the dimensionality of the problem. Finally, the relationship between the primal and dual optimal solutions is not clear to see after having changed the dual variables. In the following, we further develop model (2.39) to get a new dual-to-primal transformation in terms of $\boldsymbol{\beta}^*$ directly.

The primal solution $\mathbf{x}$ is a vector whose components are the coefficients of the cubic L_1 spline p_i, q_i, u_i and v_i. As have discussed previously, at the primal optimal solution $\mathbf{x}^*$, $p_i = z_i$, q_i is unrestricted, and $(u_i, v_i)^T$ is a two-dimensional vector satisfying

$$u_i\left(y_3^i - {y_3^i}^*\right) + v_i\left(y_4^i - {y_4^i}^*\right) \leq 0, \;\; \forall \left(y_3^i, y_4^i\right)^T \in \Theta_i, \tag{2.40}$$

i.e., $(u_i, v_i)^T$ is a normal vector of Θ_i at $\left({y_3^i}^*, {y_4^i}^*\right)^T$, where ${y_3^i}^*$ and ${y_4^i}^*$ are two components of the dual optimal solution $\mathbf{y}^*$, and $\left({y_3^i}^*, {y_4^i}^*\right)^T \in \Theta_i$. The set Θ_i is defined by

$$\begin{aligned}\Theta_i = \; & \left\{ \left(y_3^i, y_4^i\right)^T \in R^2 \middle| -h_i \le y_3^i \le h_i, \right. \\ & \left. -\tfrac{h_i^2}{2} + \tfrac{1}{4}\left(y_3^i + h_i\right)^2 \le y_4^i \le \tfrac{h_i^2}{2} - \tfrac{1}{4}\left(y_3^i - h_i\right)^2 \right\},\end{aligned} \tag{2.41}$$

which is a convex set bounded by two quadratic curves:

$$C_1^i : \; y_4^i = -\frac{h_i^2}{2} + \frac{1}{4}\left(y_3^i + h_i\right)^2,$$

and

$$C_2^i : \; y_4^i = \frac{h_i^2}{2} - \frac{1}{4}\left(y_3^i - h_i\right)^2.$$

From (2.25), (2.26), (2.27), (2.28) and (2.30), after the dual optimal solution $\boldsymbol{\beta}^*$ is obtained, another set of dual optimal solution $\mathbf{y}^*$ is calculated through an affine transformation

$$\begin{pmatrix} y_1^0 \\ y_2^0 \\ y_3^0 \\ y_4^0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{h_0} \\ 0 \\ \frac{h_0}{2} \\ \frac{h_0^2}{3} \end{pmatrix} \beta_0,$$

$$\begin{pmatrix} y_1^1 \\ y_2^1 \\ y_3^1 \\ y_4^1 \end{pmatrix} = \begin{pmatrix} \frac{1}{h_0} + \frac{1}{h_1} & -\frac{1}{h_1} \\ 0 & 0 \\ \frac{h_1}{2} & \frac{h_1}{2} \\ \frac{h_1^2}{6} & \frac{h_1^2}{3} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix},$$

$$\begin{pmatrix} y_1^i \\ y_2^i \\ y_3^i \\ y_4^i \end{pmatrix} = \begin{pmatrix} -\frac{1}{h_{i-1}} & \frac{1}{h_{i-1}} + \frac{1}{h_i} & -\frac{1}{h_i} \\ 0 & 0 & 0 \\ 0 & \frac{h_i}{2} & \frac{h_i}{2} \\ 0 & \frac{h_i^2}{6} & \frac{h_i^2}{3} \end{pmatrix} \begin{pmatrix} \beta_{i-2} \\ \beta_{i-1} \\ \beta_i \end{pmatrix}, \quad i = 2, \ldots, n-1,$$

$$\begin{pmatrix} y_1^n \\ y_2^n \end{pmatrix} = \begin{pmatrix} -\frac{1}{h_{n-1}} & \frac{1}{h_{n-1}} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{n-2} \\ \beta_{n-1} \end{pmatrix}.$$

Here we only concern those part of the affine transformation involving y_3^i and y_4^i,

$$\begin{pmatrix} y_3^0 \\ y_4^0 \end{pmatrix} = \begin{pmatrix} \frac{h_0}{2}\beta_0 \\ \frac{h_0^2}{3}\beta_0 \end{pmatrix}, \tag{2.42}$$

$$\begin{pmatrix} y_3^i \\ \\ y_4^i \end{pmatrix} = \begin{pmatrix} \frac{h_i}{2} & \frac{h_i}{2} \\ \\ \frac{h_i^2}{6} & \frac{h_i^2}{3} \end{pmatrix} \begin{pmatrix} \beta_{i-1} \\ \\ \beta_i \end{pmatrix} = H \begin{pmatrix} \beta_{i-1} \\ \\ \beta_i \end{pmatrix}, \; i = 1, \ldots, n-1. \tag{2.43}$$

For $i = 1, \ldots, n-2$, $(\beta_{i-1}, \beta_i)^T$ is defined over the set

$$\Omega = \left\{ (x, y)^T \in R^2 \, | g(x, y) \leq 0 \right\},$$

which is a convex set enclosed by two quadratic curves

$$C_L = \left\{ (x, y)^T \in R^2 \middle| -\frac{1}{3}(x - y) = 1 - \frac{1}{4}(x + y)^2 \right\},$$

and

$$C_R = \left\{ (x, y)^T \in R^2 \middle| \frac{1}{3}(x - y) = 1 - \frac{1}{4}(x + y)^2 \right\}.$$

Notice that the transformation from $(\beta_{i-1}, \beta_i)^T$ to $\left(y_3^i, y_4^i\right)^T$ defined in (2.43) is a one-to-one and onto transformation from Ω to Θ_i. $(\beta_{i-1}, \beta_i)^T$ is an interior point of Ω if and only if $\left(y_3^i, y_4^i\right)^T$ is an interior point of Θ_i. $(\beta_{i-1}, \beta_i)^T$ is on the boundary curves C_L of Ω if and only if $\left(y_3^i, y_4^i\right)^T$ is on the boundary curve C_2^i of Θ_i. $(\beta_{i-1}, \beta_i)^T$ is on the boundary curve C_R if and only if $\left(y_3^i, y_4^i\right)^T$ is on the boundary curve C_1^i of Θ_i. Following (2.40) and (2.43), we have

$$\begin{aligned}
& u_i \left(y_3^i - {y_3^i}^*\right) + v_i \left(y_4^i - {y_4^i}^*\right) \\
= \; & (u_i, v_i) \begin{pmatrix} y_3^i - {y_3^i}^* \\ \\ y_4^i - {y_4^i}^* \end{pmatrix} \\
= \; & (u_i, v_i) \, H \begin{pmatrix} \beta_{i-1} - \beta_{i-1}^* \\ \\ \beta_i - \beta_i^* \end{pmatrix} \\
= \; & \left(H^T \begin{pmatrix} u_i \\ \\ v_i \end{pmatrix} \right)^T \begin{pmatrix} \beta_{i-1} - \beta_{i-1}^* \\ \\ \beta_i - \beta_i^* \end{pmatrix} \leq 0.
\end{aligned} \tag{2.44}$$

By (2.9), (2.10), and (2.43),

$$H^T \begin{pmatrix} u_i \\ \\ v_i \end{pmatrix} = \begin{pmatrix} \frac{h_i}{2} & \frac{h_i^2}{6} \\ \\ \frac{h_i}{2} & \frac{h_i^2}{3} \end{pmatrix} \begin{pmatrix} -\frac{2}{h_i}\left(2q_i + q_{i+1} - 3\Delta z_i\right) \\ \\ \frac{6}{h_i^2}\left(q_i + q_{i+1} - 2\Delta z_i\right) \end{pmatrix} = \begin{pmatrix} -q_i + \Delta z_i \\ \\ q_{i+1} - \Delta z_i \end{pmatrix}.$$

Therefore, for $i = 1, \ldots, n-2$, $(-q_i + \Delta z_i, q_{i+1} - \Delta z_i)^T$ is a normal vector of Ω at point $\left(\beta_{i-1}^*, \beta_i^*\right)^T$, where β_{i-1}^* and β_i^* are the $(i-1)$th and ith elements of the dual optimal solution $\boldsymbol{\beta}^*$. Let $(n_1(x, y), n_2(x, y))^T$

be the normal vector of Ω at point $(x,y)^T$, then the primal optimal solution is characterized by $\boldsymbol{\beta}^*$ as

$$\begin{pmatrix} -q_i + \Delta z_i \\ \\ q_{i+1} - \Delta z_i \end{pmatrix} = \begin{pmatrix} n_1\left(\beta_{i-1}^*, \beta_i^*\right) \\ \\ n_2\left(\beta_{i-1}^*, \beta_i^*\right) \end{pmatrix}. \tag{2.45}$$

The next task is to express the normal vector $\left(n_1\left(\beta_{i-1}^*, \beta_i^*\right), n_2\left(\beta_{i-1}^*, \beta_i^*\right)\right)^T$, $i = 1, \ldots, n-2$, in terms of β_{i-1}^* and β_i^* explicitly.

If $\left(\beta_{i-1}^*, \beta_i^*\right)^T$ is an interior point of Ω, then obviously $\left(n_1\left(\beta_{i-1}^*, \beta_i^*\right), n_2\left(\beta_{i-1}^*, \beta_i^*\right)\right)^T = (0,0)^T$. Assume $\left(\beta_{i-1}^*, \beta_i^*\right)^T$ is a boundary point of Ω on the curve C_L but not on the curve C_R, then the corresponding $\left({y_3^i}^*, {y_4^i}^*\right)^T$ must be a boundary point of Θ_i on the curve C_2^i but not on the curve C_1^i. In this case, it has been shown in [2] that the normal cone of Θ_i at $\left({y_3^i}^*, {y_4^i}^*\right)^T$ is spanned by $\left(\frac{{y_3^i}^* - h_i}{2}, 1\right)^T$. Hence, by comparing (2.40) and (2.44), we have

$$\begin{aligned}
\begin{pmatrix} n_1\left(\beta_{i-1}^*, \beta_i^*\right) \\ \\ n_2\left(\beta_{i-1}^*, \beta_i^*\right) \end{pmatrix} &= \lambda_i H^T \begin{pmatrix} \frac{{y_3^i}^* - h_i}{2} \\ \\ 1 \end{pmatrix} \\
&= \lambda_i \begin{pmatrix} \frac{h_i}{2} & \frac{h_i^2}{6} \\ \\ \frac{h_i}{2} & \frac{h_i^2}{3} \end{pmatrix} \begin{pmatrix} \frac{(\beta_{i-1}^* + \beta_i^* - 2)h_i}{4} \\ \\ 1 \end{pmatrix} \\
&= \lambda_i \frac{h_i^2}{8} \begin{pmatrix} \beta_{i-1}^* + \beta_i^* - \frac{2}{3} \\ \\ \beta_{i-1}^* + \beta_i^* + \frac{2}{3} \end{pmatrix}, \quad \lambda_i \geq 0.
\end{aligned}$$

For simplicity, we write

$$\begin{pmatrix} n_1\left(\beta_{i-1}^*, \beta_i^*\right) \\ \\ n_2\left(\beta_{i-1}^*, \beta_i^*\right) \end{pmatrix} = \lambda_i \begin{pmatrix} \beta_{i-1}^* + \beta_i^* - \frac{2}{3} \\ \\ \beta_{i-1}^* + \beta_i^* + \frac{2}{3} \end{pmatrix}, \quad \lambda_i \geq 0.$$

Similarly, if $\left(\beta_{i-1}^*, \beta_i^*\right)^T$ is a boundary point of Ω on the curve C_R but not on the curve C_L, then

$$\begin{pmatrix} n_1\left(\beta_{i-1}^*, \beta_i^*\right) \\ \\ n_2\left(\beta_{i-1}^*, \beta_i^*\right) \end{pmatrix} = \mu_i \begin{pmatrix} \beta_{i-1}^* + \beta_i^* + \frac{2}{3} \\ \\ \beta_{i-1}^* + \beta_i^* - \frac{2}{3} \end{pmatrix}, \quad \mu_i \geq 0.$$

If $\left(\beta_{i-1}^*, \beta_i^*\right)^T$ is a boundary point of Ω on both the curve C_L and the curve C_R, then we say $\left(\beta_{i-1}^*, \beta_i^*\right)^T$ is a corner point of Ω. In this case, $\left(\beta_{i-1}^*, \beta_i^*\right)^T = (1,1)^T$ or $(-1,-1)^T$. If $\left(\beta_{i-1}^*, \beta_i^*\right)^T = (1,1)^T$, then

$$\begin{pmatrix} n_1\left(\beta_{i-1}^*, \beta_i^*\right) \\ \\ n_2\left(\beta_{i-1}^*, \beta_i^*\right) \end{pmatrix} = \lambda_i \begin{pmatrix} 1 \\ \\ 2 \end{pmatrix} + \mu_i \begin{pmatrix} 2 \\ \\ 1 \end{pmatrix}, \quad \lambda_i, \mu_i \geq 0.$$

If $\left(\beta_{i-1}^{*}, \beta_{i}^{*}\right)^{T}=(-1,-1)^{T}$, then

$$\left(\begin{array}{c} n_{1}\left(\beta_{i-1}^{*}, \beta_{i}^{*}\right) \\ \\ n_{2}\left(\beta_{i-1}^{*}, \beta_{i}^{*}\right) \end{array}\right)=\lambda_{i}\left(\begin{array}{c} -1 \\ \\ -2 \end{array}\right)+\mu_{i}\left(\begin{array}{c} -2 \\ \\ -1 \end{array}\right), \quad \lambda_{i}, \mu_{i} \geq 0.$$

For $i=0$, by the relationship (2.42), if $\left|\beta_{0}^{*}\right|<\frac{\sqrt{40}-2}{3}$, then $\left(y_{3}^{0^{*}}, y_{4}^{0^{*}}\right)^{T}$ is an interior point of Θ_{0}. It is clear that $\left(u_{0}, v_{0}\right)^{T}=(0,0)^{T}$, which implies $\left(q_{0}, q_{1}\right)^{T}=\left(\Delta z_{0}, \Delta z_{0}\right)^{T}$ by (2.9) and (2.10). If $\beta_{0}^{*}=\frac{\sqrt{40}-2}{3}$, then $\left(y_{3}^{0^{*}}, y_{4}^{0^{*}}\right)^{T}$ is a boundary point of Θ_{0} on the curve C_{2}^{0}. In this case, it has been shown before that

$$\left(\begin{array}{c} u_{0} \\ \\ v_{0} \end{array}\right)=\lambda_{0}\left(\begin{array}{c} \frac{y_{3}^{0^{*}}-h_{0}}{2} \\ \\ 1 \end{array}\right)=\lambda_{0}\left(\begin{array}{c} \frac{h_{0}}{4}\left(\beta_{0}^{*}-2\right) \\ \\ 1 \end{array}\right), \quad \lambda_{0} \geq 0.$$

Hence, by (2.9) and (2.10)

$$\left(\begin{array}{c} q_{0} \\ \\ q_{1} \end{array}\right)=\left(\begin{array}{c} \Delta z_{0}-\lambda_{0} \frac{h_{0}^{2}}{8}\left(\beta_{0}^{*}-\frac{2}{3}\right) \\ \\ \Delta z_{0}+\lambda_{0} \frac{h_{0}^{2}}{8}\left(\beta_{0}^{*}+\frac{2}{3}\right) \end{array}\right), \quad \lambda_{0} \geq 0.$$

If $\beta_{0}^{*}=-\frac{\sqrt{40}-2}{3}$, then $\left(y_{3}^{0^{*}}, y_{4}^{0^{*}}\right)^{T}$ is a boundary point of Θ_{0} on the curve C_{1}^{0}. In this case, we have shown before that

$$\left(\begin{array}{c} u_{0} \\ \\ v_{0} \end{array}\right)=\lambda_{0}\left(\begin{array}{c} \frac{y_{3}^{0^{*}}+h_{i}}{2} \\ \\ -1 \end{array}\right)=\lambda_{0}\left(\begin{array}{c} \frac{h_{0}}{4}\left(\beta_{0}^{*}+2\right) \\ \\ -1 \end{array}\right), \quad \lambda_{0} \geq 0.$$

Again, by (2.9) and (2.10)

$$\left(\begin{array}{c} q_{0} \\ \\ q_{1} \end{array}\right)=\left(\begin{array}{c} \Delta z_{0}-\lambda_{0} \frac{h_{0}^{2}}{8}\left(\beta_{0}^{*}+\frac{2}{3}\right) \\ \\ \Delta z_{0}+\lambda_{0} \frac{h_{0}^{2}}{8}\left(\beta_{0}^{*}-\frac{2}{3}\right) \end{array}\right), \quad \lambda_{0} \geq 0.$$

In summary, for $i=0$, we have

$$\left(\begin{array}{c} q_{0} \\ \\ q_{1} \end{array}\right)=\left(\begin{array}{c} \Delta z_{0}-n_{1}\left(\beta_{0}^{*}\right) \\ \\ \Delta z_{0}+n_{2}\left(\beta_{0}^{*}\right) \end{array}\right), \tag{2.46}$$

where

$$\begin{pmatrix} n_1\left(\beta_0^*\right) \\ n_2\left(\beta_0^*\right) \end{pmatrix} = \begin{cases} \lambda_0 \begin{pmatrix} \beta_0^* - \frac{2}{3} \\ \beta_0^* + \frac{2}{3} \end{pmatrix}, & \lambda_0 \geq 0, \quad \text{if } \beta_0^* = \frac{\sqrt{40}-2}{3}, \\ \lambda_0 \begin{pmatrix} \beta_0^* + \frac{2}{3} \\ \beta_0^* - \frac{2}{3} \end{pmatrix}, & \lambda_0 \geq 0, \quad \text{if } \beta_0^* = -\frac{\sqrt{40}-2}{3}, \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & \text{otherwise.} \end{cases} \tag{2.47}$$

For $i = n-1$, notice that $\beta_{n-1}^* = 0$. Then, from (2.43),

$$\begin{pmatrix} y_3^{n-1*} \\ y_4^{n-1*} \end{pmatrix} = \begin{pmatrix} \frac{h_{n-1}}{2}\beta_{n-2}^* \\ \frac{h_{n-1}^2}{6}\beta_{n-2}^* \end{pmatrix}.$$

Following the same logic as for the case $i = 0$, we can derive that

$$\begin{pmatrix} q_{n-1} \\ q_n \end{pmatrix} = \begin{pmatrix} \Delta z_{n-1} - n_1\left(\beta_{n-2}^*\right) \\ \Delta z_{n-1} + n_2\left(\beta_{n-2}^*\right) \end{pmatrix}, \tag{2.48}$$

where

$$\begin{pmatrix} n_1\left(\beta_{n-2}^*\right) \\ n_2\left(\beta_{n-2}^*\right) \end{pmatrix} = \begin{cases} \lambda_{n-1} \begin{pmatrix} \beta_{n-2}^* + \frac{2}{3} \\ \beta_{n-2}^* - \frac{2}{3} \end{pmatrix}, & \lambda_{n-1} \geq 0, \quad \text{if } \beta_{n-2}^* = \frac{\sqrt{40}-2}{3}, \\ \lambda_{n-1} \begin{pmatrix} \beta_{n-2}^* - \frac{2}{3} \\ \beta_{n-2}^* + \frac{2}{3} \end{pmatrix}, & \lambda_{n-1} \geq 0, \quad \text{if } \beta_{n-2}^* = -\frac{\sqrt{40}-2}{3}, \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & \text{otherwise.} \end{cases} \tag{2.49}$$

Therefore, once the dual optimal solution $\boldsymbol{\beta}^*$ is obtained, the primal optimal solution can be obtained by solving

$$
\text{(T)} \quad \begin{cases}
\min\limits_{q_i, \lambda_i, \mu_i} & \sum_{i=0}^{n} |q_i| \\
\text{s.t.} & q_0 = \Delta z_0 - n_1\left(\beta_0^*\right), \\
& q_1 = \Delta z_0 + n_2\left(\beta_0^*\right), \\
& q_i = \Delta z_i - n_1\left(\beta_{i-1}^*, \beta_i^*\right),\ i = 1, \ldots, n-2, \\
& q_{i+1} = \Delta z_i + n_2\left(\beta_{i-1}^*, \beta_i^*\right),\ i = 1, \ldots, n-2, \\
& q_{n-1} = \Delta z_{n-1} - n_1\left(\beta_{n-2}^*\right), \\
& q_n = \Delta z_{n-1} + n_2\left(\beta_{n-2}^*\right), \\
& \lambda_i, \mu_i \geq 0,\ i = 0, \ldots, n-1.
\end{cases}
$$

The generalized geometric programming theory guarantees that any feasible solution of the problem (T) results in a cubic L_1 spline. However, the feasible set of (T) contains more than one single point in most of the case. Hence, we need a criterion to pick one from all candidate splines. This criterion is reflected in the objective function of (T), which is called a regularization function. In the geometric design society, people usually prefers the flattest curve, i.e., the curve with the smallest absolute values of q_i's. Therefore, the regularization function is chosen to be the l_p norm of $\mathbf{q}$. Among them, the l_1 norm of $\mathbf{q}$ enables the problem (T) to become a linear program. Consequently, $\sum_{i=0}^{n} |q_i|$ is the most frequently used regularization function.

2.4 Conclusion

We have derived a mathematical model for the univariate cubic L_1 spline problem and have shown that solving a univariate cubic L_1 spline problem is equivalent to solving an unconstrained optimization problem. The objective function of this optimization problem is convex and almost everywhere differentiable.

The generalized geometric programming formulation provides an attractive theoretical framework for calculating the coefficients of univariate cubic L_1 splines. We have explicitly constructed a geometric dual problem and corresponding optimality conditions. Compared to the discretization approach, which uses an interior point algorithm [7] to solve the primal problem, the geometric programming approach has a clear conceptual advantage in that it is based on minimization of the exact continuum spline functional and does not depend on numerical approximation of this functional, which could lead to different users producing different L_1 splines for the same data.

The discretization approach has already been applied in bivariate situations [16] and could be applied in multivariate situations. However, significantly different L_1 splines may result from different discretizations. Moreover, interior-point methods for solving the large sets of linear systems that result from the numerical discretization of bi/multivariate integrals will be computationally very expensive. After developing a specialized geometric programming algorithm for the univariate case, we hope to extend this algorithm to bi- and multivariate cases, which are of much greater interest in the geometric modeling community than the univariate case. The univariate theory developed in the present paper is one step on the long path toward a new class of optimization procedures that will, we hope and expect, be the basis for greatly improved geometric modeling by L_1 splines.

2.5 Calculation of conjugate transform $\mathfrak{h}(\mathbf{y}) : \mathfrak{D}$

From the definition of conjugate transform, we have

$$\begin{aligned}
&\mathfrak{h}_i\left(\mathbf{y}^i\right) \\
&= \sup_{\mathbf{x}^i \in \mathfrak{C}_i} \left[\left\langle \mathbf{y}^i, \mathbf{x}^i \right\rangle - \mathfrak{g}_i\left(\mathbf{x}^i\right)\right] \\
&= \sup_{p_i = z_i} \left[y_1^i p_i\right] + \sup_{q_i \in R} \left[y_2^i q_i\right] + \sup_{u_i, v_i \in R} \left[y_3^i u_i + y_4^i v_i - \mathfrak{g}_i\left(\mathbf{x}^i\right)\right]. \qquad (2.50)
\end{aligned}$$

In order to make sure that $\mathfrak{h}_i\left(\mathbf{y}^i\right) < +\infty$, each of the three terms in (2.50) should be finite. Since $p_i = z_i$ is a fixed real number, the first term is always finite and the supremum is given by $z_i y_1^i$. For the second term, since q_i can be any real number, the only value of y_2^i which makes this term finite is 0, and $\sup_{q_i \in R}\left[y_2^i q_i\right] = 0$. Consequently,

$$\mathfrak{h}_i\left(\mathbf{y}^i\right) = z_i y_1^i + \sup_{u_i, v_i \in R} \left[y_3^i u_i + y_4^i v_i - \mathfrak{g}_i\left(\mathbf{x}^i\right)\right],$$

and

$$y_1^i \in R, \quad y_2^i = \{0\}.$$

The most difficult is the last term. In order to facilitate the discussion, we simplify the notation. Denote

$$\begin{aligned}&\mathfrak{f}\left(u,v\right)\\&=\begin{cases}\frac{h}{2}\left[\left|\frac{h}{2}v\right|+\frac{\left(u+\frac{h}{2}v\right)^2}{\left|\frac{h}{2}v\right|}\right], & \text{if } \left|\frac{h}{2}v\right|>\left|u+\frac{h}{2}v\right|,\\ h\left|u+\frac{h}{2}v\right|, & \text{if } \left|\frac{h}{2}v\right|\leq\left|u+\frac{h}{2}v\right|,\end{cases}\end{aligned}$$

where $(u,v)\in R^2$ and $h>0$ is a constant real number. Let the conjugate transform of $\mathfrak{f}\left(u,v\right):R^2$ be $\omega\left(\xi,\eta\right):\Omega$, where

$$\Omega=\left\{(\xi,\eta)\in R^2\,\middle|\,\sup_{(u,v)\in R^2}\left[\xi u+\eta v-\mathfrak{f}\left(u,v\right)\right]<+\infty\right\},$$

and

$$\omega\left(\xi,\eta\right)=\sup_{(u,v)\in R^2}\left[\xi u+\eta v-\mathfrak{f}\left(u,v\right)\right],\quad\forall\left(\xi,\eta\right)\in\Omega.$$

Notice that the two-dimensional (u,v)-plane is divided into four disjoint parts by the lines given by $\left|\frac{h}{2}v\right|=\left|u+\frac{h}{2}v\right|$. Let Ω_k, $k=1,2,3$, and 4, be the sets on which the function $\xi u+\eta v-\mathfrak{f}\left(u,v\right)$ is finite on each of these four parts of the (u,v)-plane and let

$$\omega_k\left(\xi,\eta\right)=\sup_{(u,v)\in R^2}\left[\xi u+\eta v-\mathfrak{f}\left(u,v\right)\right],\quad\forall\left(\xi,\eta\right)\in\Omega_k.$$

Clearly, the domain Ω of ω is the intersection of all the Ω_k and the conjugate transform $\omega\left(\xi,\eta\right)$ can be obtained as the maximum of $\omega_k\left(\xi,\eta\right)$, $k=1,2,3,4$. We treat each of the four partitions of the (u,v)-plane separately.

- **Case 1:** $u+\frac{h}{2}v\geq 0$ and $\frac{h}{2}v\geq 0$.

In this case, when $\frac{h}{2}v\leq u+\frac{h}{2}v$, we have

$$u\geq 0\ \text{ and }\ v\geq 0,$$

and

$$\mathfrak{f}\left(u,v\right)=hu+\frac{h^2}{2}v.$$

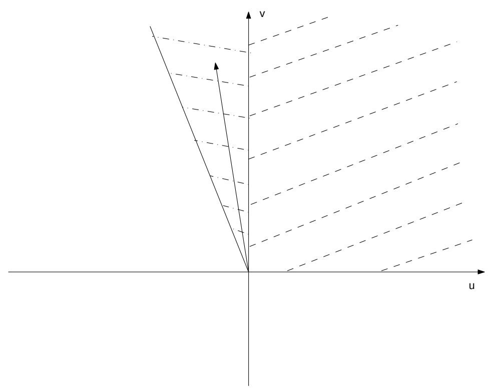

Figure 2.3: Calculate the conjugate transform of $\mathfrak{f}\,(u,v)$: case 1.

Consequently,

$$\begin{aligned}
&\omega_{1a}(\xi,\eta)\\
&=\sup_{u\geq 0, v\geq 0}\left[\xi u+\eta v-\mathfrak{f}\,(u,v)\right]\\
&=\sup_{u\geq 0, v\geq 0}\left[\xi u+\eta v-hu-\frac{h^2}{2}v\right]\\
&=\sup_{u\geq 0, v\geq 0}\left[(\xi-h)\;u+\left(\eta-\frac{h^2}{2}\right)\;v\right]\\
&=0.
\end{aligned}$$

The set on which $\omega_{1a}(\xi,\eta)$ is finite is given as

$$\Omega_{1a}=\left\{(\xi,\eta)\in R^2\,\middle|\,\xi\leq h,\ \eta\leq\frac{h^2}{2}\right\}.$$

When $\frac{h}{2}v>u+\frac{h}{2}v$, we have

$$0\geq u\geq -\frac{h}{2}v\quad\text{and}\quad v\geq 0,$$

and

$$\mathfrak{f}\,(u,v)=\frac{h}{2}\left[\frac{h}{2}v+\frac{\left(u+\frac{h}{2}v\right)^2}{\frac{h}{2}v}\right].$$

Let

$$t=\frac{u}{\frac{h}{2}v},\quad t\in[-1,0].$$

Then

$$u = t \cdot \frac{h}{2} v, \quad t \in [-1, 0], \quad v \geq 0,$$

and

$$\begin{aligned} \mathfrak{f}(u, v) &= \frac{h}{2}\left[\frac{h}{2} v + \frac{\left(u + \frac{h}{2} v\right)^2}{\frac{h}{2} v}\right] \\ &= \left(t^2 + 2t + 2\right) \cdot \frac{h^2}{4} \cdot v. \end{aligned}$$

Hence,

$$\begin{aligned} &\omega_{1b}(\xi, \eta) \\ &= \sup_{0 \geq u \geq -\frac{h}{2} v, v \geq 0} \left[\xi u + \eta v - \mathfrak{f}(u, v)\right] \\ &= \sup_{v \geq 0, t \in [-1,0]} \left[\xi t \cdot \frac{h}{2} v + \eta v - \left(t^2 + 2t + 2\right) \cdot \frac{h^2}{4} \cdot v\right] \\ &= \sup_{v \geq 0, t \in [-1,0]} \left[t^2 + 2\left(1 - \frac{\xi}{h}\right) t + 2\left(1 - \frac{2\eta}{h^2}\right)\right]\left(-\frac{h^2}{4} v\right) \\ &= 0. \end{aligned}$$

The set on which $\omega_{1b}(\xi, \eta)$ is finite is given by

$$\begin{aligned} \Omega_{1b} &= \left\{(\xi, \eta) \in R^2 \,\middle|\, t^2 + 2\left(1 - \frac{\xi}{h}\right) t + 2\left(1 - \frac{2\eta}{h^2}\right) \geq 0, \quad \forall t \in [-1, 0]\right\} \\ &= \left\{(\xi, \eta) \in R^2 \,\middle|\, \left[t + \left(1 - \frac{\xi}{h}\right)\right]^2 - \frac{1}{h^2}\left[(\xi - h)^2 + 4\left(\eta - \frac{h^2}{2}\right)\right] \geq 0, \quad \forall t \in [-1, 0]\right\}. \end{aligned}$$

Define

$$r(t) = \left[t + \left(1 - \frac{\xi}{h}\right)\right]^2 - \frac{1}{h^2}\left[(\xi - h)^2 + 4\left(\eta - \frac{h^2}{2}\right)\right], \; t \in [-1, 0].$$

If

$$(\xi - h)^2 + 4\left(\eta - \frac{h^2}{2}\right) \leq 0,$$

then

$$r(t) \geq 0, \quad \text{for all } t \in R^1.$$

Otherwise,

$$(\xi - h)^2 + 4\left(\eta - \frac{h^2}{2}\right) \geq 0.$$

Under this condition, the two roots of $r(t)$ are given by

$$\widehat{t}_1 = -\left(1-\frac{\xi}{h}\right) - \sqrt{\left(1-\frac{\xi}{h}\right)^2 + 4\left(\frac{\eta}{h^2}-\frac{1}{2}\right)},$$

and

$$\widehat{t}_2 = -\left(1-\frac{\xi}{h}\right) + \sqrt{\left(1-\frac{\xi}{h}\right)^2 + 4\left(\frac{\eta}{h^2}-\frac{1}{2}\right)}.$$

Therefore, $r(t) \geq 0$, $\forall\, t \in [-1, 0]$, if and only if

$$\widehat{t}_1 \geq 0 \ \text{ or } \ \widehat{t}_2 \leq -1.$$

Notice that $\widehat{t}_1 \geq 0$ is equivalent to

$$\xi \geq h,\ \eta \leq \frac{h^2}{2},$$

and $\widehat{t}_2 \leq -1$ means that

$$\xi \leq 0,\ \eta \leq \frac{h}{2}\xi + \frac{h^2}{4}.$$

Defining

$$\Omega_{1b}^1 = \left\{ (\xi,\eta) \in R^2 \,\middle|\, (\xi-h)^2 + 4\left(\eta - \frac{h^2}{2}\right) \leq 0 \right\},$$

$$\begin{aligned}\Omega_{1b}^2 = \ & \Big\{ (\xi,\eta) \in R^2 \,\Big|\, (\xi-h)^2 + 4\left(\eta - \tfrac{h^2}{2}\right) \geq 0, \\ & \quad \text{and } \ \xi \geq h,\ \eta \leq \tfrac{h^2}{2} \Big\},\end{aligned}$$

and

$$\begin{aligned}\Omega_{1b}^3 = \ & \Big\{ (\xi,\eta) \in R^2 \,\Big|\, (\xi-h)^2 + 4\left(\eta - \tfrac{h^2}{2}\right) \geq 0, \\ & \quad \text{and } \ \xi \leq 0,\ \eta \leq \tfrac{h}{2}\xi + \tfrac{h^2}{4} \Big\},\end{aligned}$$

the set on which $\omega_{1b}(\xi,\eta)$ is finite is given by

$$\Omega_{1b} = \Omega_{1b}^1 \cup \Omega_{1b}^2 \cup \Omega_{1b}^3.$$

Therefore, the domain of the conjugate transform for case 1 is defined as

$$\begin{aligned}\Omega_1 &= \left\{ (\xi,\eta) \in R^2 \,\middle|\, \sup_{u+\frac{h}{2}v \geq 0,\, \frac{h}{2}v \geq 0} \left[\xi u + \eta v - \mathfrak{f}(u,v)\right] < +\infty \right\} \\ &= \Omega_{1a} \cap \Omega_{1b} = \Omega_{1a} \cap \left[\Omega_{1b}^1 \cup \Omega_{1b}^2 \cup \Omega_{1b}^3\right] \\ &= \begin{array}{l} \Big\{ (\xi,\eta) \in R^2 \,\Big|\, 0 \leq \xi \leq h,\ \eta \leq \tfrac{h^2}{2} - \tfrac{1}{4}(\xi-h)^2 \\ \qquad \text{or } \ \xi \leq 0,\ \eta \leq \tfrac{h}{2}\xi + \tfrac{h^2}{4} \Big\} \end{array}.\end{aligned}$$

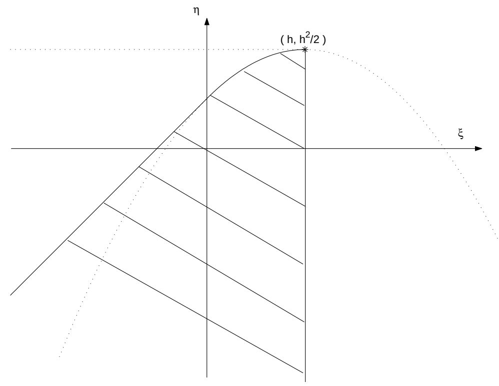

Figure 2.4: Domain of the conjugate transform: case 1.

Ω_1 is the shaded area in Figure 2.4.

Now, for any $(\xi, \eta) \in \Omega_1$,

$$\begin{aligned} &\omega_1(\xi, \eta) \\ &= \max\{\omega_{1a}(\xi, \eta), \omega_{1b}(\xi, \eta)\} \\ &= 0. \end{aligned}$$

For the other three cases, the calculations are similiar. We mention only the following items of interest:

- **Case 2:** $u + \frac{h}{2}v \leq 0$ and $\frac{h}{2}v \geq 0$.

The domain of the conjugate transform for case 2 is defined as

$$\Omega_2 = \begin{array}{c} \left\{(\xi, \eta) \in R^2 \middle| -h \leq \xi \leq 0,\ \eta \leq \frac{h^2}{2} - \frac{1}{4}(\xi - h)^2 \right. \\ \left. \text{or}\ \ \xi \geq 0,\ \eta \leq \frac{h}{2}\xi + \frac{h^2}{4} \right\}. \end{array}$$

For any $(\xi, \eta) \in \Omega_2$,

$$\omega_2(\xi, \eta) = 0.$$

- **Case 3:** $u + \frac{h}{2}v \leq 0$ and $\frac{h}{2}v \leq 0$.

The domain of the conjugate transform for case 3 is defined as

$$\Omega_3 = \begin{array}{c} \left\{ (\xi,\eta) \in R^2 \middle| -h \le \xi \le 0,\ \eta \ge -\frac{h^2}{2} + \frac{1}{4}(\xi+h)^2 \right. \\ \left. \text{or} \ \ \xi \ge 0,\ \eta \ge \frac{h}{2}\xi - \frac{h^2}{4} \right\}. \end{array}$$

For any $(\xi,\eta) \in \Omega_3$,

$$\omega_3(\xi,\eta) = 0.$$

- **Case 4:** $u + \frac{h}{2}v \ge 0$ and $\frac{h}{2}v \le 0$.

The domain of the conjugate transform for this case is defined as

$$\Omega_4 = \begin{array}{c} \left\{ (\xi,\eta) \in R^2 \middle| 0 \le \xi \le h,\ \eta \ge -\frac{h^2}{2} + \frac{1}{4}(\xi+h)^2 \right. \\ \left. \text{or} \ \ \xi \le 0,\ \eta \ge \frac{h}{2}\xi - \frac{h^2}{4} \right\}. \end{array}$$

For any $(\xi,\eta) \in \Omega_4$,

$$\omega_4(\xi,\eta) = 0.$$

Any $(u,v) \in R^2$ falls into one of the above four cases. In order to guarantee that $\xi u + \eta v - \mathfrak{f}(u,v)$ is finite for any (u,v), we need to require that $\xi u + \eta v - \mathfrak{f}(u,v)$ is finite in each of these four cases. This means

$$\begin{aligned} \Omega &= \left\{ (\xi,\eta) \in R^2 \middle| \sup_{(u,v)\in R^2} [\xi u + \eta v - \mathfrak{f}(u,v)] < +\infty \right\} \\ &= \Omega_1 \cap \Omega_2 \cap \Omega_3 \cap \Omega_4. \end{aligned}$$

Notice that the function $\eta = \frac{h^2}{2} - \frac{1}{4}(\xi - h)^2$ is a concave function and, at the point $(\xi,\eta) = \left(0, \frac{h^2}{4}\right)$, its tangent is given by $\eta = \frac{h}{2}\xi + \frac{h^2}{4}$. Similarly, the function $\eta = -\frac{h^2}{2} + \frac{1}{4}(\xi+h)^2$ is a convex function and, at the point $(\xi,\eta) = \left(0, -\frac{h^2}{4}\right)$, its tangent is given by $\eta = \frac{h}{2}\xi - \frac{h^2}{4}$. Moreover, these two quadratic functions $\eta = \frac{h^2}{2} - \frac{1}{4}(\xi-h)^2$ and $\eta = -\frac{h^2}{2} + \frac{1}{4}(\xi+h)^2$ intersect at two points $\left(h, \frac{h^2}{2}\right)$ and $\left(-h, -\frac{h^2}{2}\right)$. Considering all these properties at the same time, the domain of $\omega(\xi,\eta)$ becomes

$$\begin{aligned} \Omega &= \Omega_1 \cap \Omega_2 \cap \Omega_3 \cap \Omega_4 \\ &= \begin{array}{c} \left\{ (\xi,\eta) \in R^2 \quad | -h \le \xi \le h, \right. \\ \left. -\frac{h^2}{2} + \frac{1}{4}(\xi+h)^2 \le \eta \le \frac{h^2}{2} - \frac{1}{4}(\xi-h)^2 \right\}. \end{array} \end{aligned}$$

Furthermore, for any $(\xi, \eta) \in \Omega$,

$$\begin{aligned}
\omega(\xi, \eta) &= \sup_{(u,v) \in R^2} [\xi u + \eta v - \mathfrak{f}(u, v)] \\
&= \max_{i=1}^{4} \{\omega_i(\xi, \eta)\} \\
&= 0.
\end{aligned}$$

Therefore the conjugate transform $\mathfrak{h}_i(\mathbf{y}^i) : \mathfrak{D}_i$, where $\mathbf{y}^i = (y_1^i, y_2^i, y_3^i, y_4^i)$, can be expressed as:

$$\mathfrak{D}_i = R \times \{0\} \times \Theta_i$$

with

$$\begin{aligned}
\Theta_i = \quad & \Big\{ (y_3^i, y_4^i) \in R^2 \Big| - h_i \leq y_3^i \leq h_i, \\
& -\tfrac{h_i^2}{2} + \tfrac{1}{4} (y_3^i + h_i)^2 \leq y_4^i \leq \tfrac{h_i^2}{2} - \tfrac{1}{4} (y_3^i - h_i)^2 \Big\}.
\end{aligned}$$

For any $\mathbf{y}^i \in \mathfrak{D}_i$,

$$\begin{aligned}
\mathfrak{h}_i(\mathbf{y}^i) &= z_i y_1^i + \sup_{u_i, v_i \in \Omega_i} [y_3^i u_i + y_4^i v_i - \mathfrak{g}_i(\mathbf{x}^i)] \\
&= z_i y_1^i.
\end{aligned}$$

Chapter 3

A geometric programming framework for univariate cubic L_1 smoothing splines

Given a set of data, our objective is to construct smooth curves or surfaces, called smoothing splines, to approximate the data. Many variants of smoothing splines have been used in science and engineering [8, 12]. Closed related with cubic L_1 interpolating splines, a class of cubic L_1 smoothing splines is introduced in [15], which are calculated by minimizing a convex combination of the l_1 norm of the residuals of the data-fitting equation and the L_1 norm of the second derivative of the splines over a finite-dimensional spline space. Numerical experiments have shown that this new class of smoothing splines perform well on preserving the shapes for arbitrary data.

Noting the similarity between interpolating splines and smoothing splines, we follow the previous chapter to formulate the problem in the context of generalized geometric programming theory [6, 21, 22, 23, 26]. We will show that the geometric dual of the smoothing-spline problem has a linear objective function over a convex feasible domain, despite the fact that the objective function of the primal problem is nonlinear and non-differentiable. In this framework, it becomes possible to perform sensitivity analysis and to develop continuum-based computational algorithms.

This chapter is organized as follows: In Section 3.1, the mathematical structure of univariate cubic L_1 smoothing splines is introduced. In Section 3.2, the problem is reformulated as a geometric programming problem. In Section 3.3, some computational results are reported. Sensitivity analysis for data with uncertainty is performed in Section 3.4. Discussion and the conclusion are in Section 3.5.

3.1 Univariate cubic L_1 smoothing splines

Given a set of data $(\hat{x}_m, \hat{z}_m)$, $m = 1, 2, \ldots, M$, with the abscissae $\hat{x}_m$ contained in a finite real interval $[a, b]$, we partition the interval by using a strictly monotonic set $\Delta = \{x_i\}_0^n$ such that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\,, \tag{3.1}$$

into n subintervals, with $I_i = [x_i, x_{i+1})$ for $i = 0, 1, \ldots, n-2$, and $I_{n-1} = [x_{n-1}, x_n]$. To find an approximation function $S(x)$ in the space of C^1 smooth, piecewise cubic polynomials over these subintervals, we characterize $S(x)$ by its function value $S(x_i)$ and the value of its first derivative $\frac{dS}{dx}(x_i)$ at the knots x_i for $i = 1, \ldots, n$. To determine these values, [15] proposed minimizing the following functional:

$$E = \lambda \sum_{m=1}^{M} \hat{w}_m \, \left|S(\hat{x}_m) - \hat{z}_m\right| + (1-\lambda) \int_{x_0}^{x_n} \left|S''(x)\right| dx, \tag{3.2}$$

where $\lambda \in (0, 1)$ and the weights $\hat{w}_m > 0$ are predetermined.

Functional E is a convex combination of an approximation error and an "energy integral". The sum is the overall weighted approximation error with a positive weight $\hat{w}_m$ at each data location $\hat{x}_m$. It measures how close the approximation function can fit the data. The integral term is one type of the bending energy that defines a cubic L_1 interpolating spline [14]. The parameter λ is a balance parameter that adjusts the trade-off between these two terms. In this paper, we consider the unconstrained minimization of E. Boundary conditions and/or other conditions on the coefficients of the smoothing splines could be imposed without significant changes in the theory and computational methods presented below.

A cubic L_1 smoothing spline is a C^1 piecewise cubic polynomial that minimizes (3.2). On each subinterval $[x_i, x_{i+1}]$, there exists a cubic polynomial $S_i(x)$,

$$S_i(x) = p_i + q_i(x - x_i) + \frac{u_i}{2}(x - x_i)^2 + \frac{v_i}{6}(x - x_i)^3, \tag{3.3}$$

such that $S(x) = S_i(x)$, $\forall x \in [x_i, x_{i+1}]$, $i = 0, \ldots, n-1$.

For each $i = 0, \ldots, n-1$, let

$$h_i = x_{i+1} - x_i > 0. \tag{3.4}$$

On the subinterval $[x_i, x_{i+1}]$, the second derivative of $S_i(x)$ is given by

$$S_i''(x) = u_i + v_i(x - x_i).$$

Changing the variable from x to $t = (x - x_i - \frac{h_i}{2})/h_i$, $-\frac{1}{2} \le t \le \frac{1}{2}$, we obtain

$$S_i''(t) = u_i + \frac{h_i}{2}v_i + h_i v_i t.$$

Then the integral in the minimization functional can be expressed explicitly as

$$\begin{aligned}
&\int_{x_0}^{x_n} |S''(x)|\, dx \\
&= \sum_{i=0}^{n-1} h_i \int_{-1/2}^{1/2} \left| u_i + \frac{h_i}{2}v_i + h_i v_i t \right| dt \qquad (3.5)\\
&= \sum_{i=0}^{n-1} \begin{cases} \frac{h_i}{2}\left[\left|\frac{h_i}{2}v_i\right| + \frac{u_i + \frac{h_i}{2}v_i}{\frac{h_i}{2}v_i}^{2} \right], & \text{if } \left|\frac{h_i}{2}v_i\right| > \left|u_i + \frac{h_i}{2}v_i\right| \\ h_i \left|u_i + \frac{h_i}{2}v_i\right|, & \text{if } \left|\frac{h_i}{2}v_i\right| \le \left|u_i + \frac{h_i}{2}v_i\right| \end{cases}. \qquad (3.6)
\end{aligned}$$

Each data location $\hat{x}_m$ falls into precisely one subinterval $[x_i, x_{i+1})$, $i = 0, 1, \ldots, n-2$ or $[x_{n-1}, x_n]$. We assume that, in each subinterval $[x_i, x_{i+1})$, $i = 0, 1, \ldots, n-2$, there are M_i data points $\hat{x}_{i1}$, …, $\hat{x}_{iM_i}$ with $x_i \le \hat{x}_{i1} < \cdots < \hat{x}_{iM_i} < x_{i+1}$, and M_{n-1} data points $\hat{x}_{n-1,1}$, …, $\hat{x}_{n-1,M_{n-1}}$ in subinterval $[x_{n-1}, x_n]$, with $x_{n-1} \le \hat{x}_{n-1,1} < \cdots < \hat{x}_{n-1,M_{n-1}} \le x_n$ such that

$$\sum_{i=0}^{n-1} M_i = M. \tag{3.7}$$

The value corresponding to the data location $\hat{x}_{ik}$ is $\hat{z}_{ik}$. Denote

$$\hat{h}_{ik} = \hat{x}_{ik} - x_i \quad \text{for } i = 0, \ldots, n-1 \quad \text{and } k = 1, \ldots M_i, \tag{3.8}$$

then $0 \le \hat{h}_{ik} \le h_i$ and the overall approximation error can be expressed explicitly in terms of the coefficients of the smoothing spline $S(x)$ as

$$\sum_{i=0}^{n-1}\sum_{k=1}^{M_i} \hat{w}_m \left| p_i + \hat{h}_{ik} q_i + \frac{\hat{h}_{ik}^2}{2} u_i + \frac{\hat{h}_{ik}^3}{6} v_i - \hat{z}_{ik} \right|. \tag{3.9}$$

The smoothing spline $S(x)$ should satisfy the C^1 smoothness constraints, that is,

$$p_i + h_i q_i + \frac{h_i^2}{2} u_i + \frac{h_i^3}{6} v_i = p_{i+1}, \qquad i = 0, \ldots, n-1, \tag{3.10}$$

$$q_i + h_i u_i + \frac{h_i^2}{2} v_i = q_{i+1}, \qquad i = 0, \ldots, n-1. \tag{3.11}$$

From constraints (3.10) and (3.11), we can derive u_i and v_i in terms of p_i and q_i, i.e.,

$$u_i = -\frac{6}{h_i^2} p_i - \frac{4}{h_i} q_i + \frac{6}{h_i^2} p_{i+1} - \frac{2}{h_i} q_{i+1}, \qquad i = 0, \ldots, n-1, \tag{3.12}$$

$$v_i = \frac{12}{h_i^3} p_i + \frac{6}{h_i^2} q_i - \frac{12}{h_i^3} p_{i+1} + \frac{6}{h_i^2} q_{i+1}, \qquad i = 0, \ldots, n-1. \tag{3.13}$$

Consequently, functional (3.2) is a function of p_i's and q_i's. Let us denote it by $E(\mathbf{p}, \mathbf{q})$, where $\mathbf{p}$ and $\mathbf{q}$ are the vectors with elements p_i and q_i, respectively. The problem of finding a smoothing spline $S(x)$ that minimizes functional (3.2) is equivalent to an unconstrained optimization problem

$$\min_{\mathbf{p},\mathbf{q} \in R^{n+1}} E(\mathbf{p}, \mathbf{q}). \tag{3.14}$$

Functional $E(\mathbf{p}, \mathbf{q})$ is convex, continuous in $\mathbf{p}$ and $\mathbf{q}$ and increases to ∞ as the Euclidean norm of $\mathbf{p}$ and $\mathbf{q}$ increases to ∞. Therefore, $E(\mathbf{p}, \mathbf{q})$ has a minimum [15]. However, $E(\mathbf{p}, \mathbf{q})$ is not strictly convex, so its minimum need not be unique. We will deal with this nonuniqueness in the dual-to-primal transformation in Subsection 3.4 below.

3.2 Generalized geometric programming approach

Generalized geometric programming is an optimization theory that provides existence, uniqueness and other theorems for mathematical analysis. A thorough treatment can be found in [21, 22]. Some new advances can be found in [26]. In this section, we show how to formulate the cubic L_1 smoothing spline problem in a generalized geometric programming setting.

3.2.1 Primal problem

For $i = 0, \ldots, n-1$ and $k = 1, \ldots M_i$, denote

$$s_{ik} = S_i\left(\hat{x}_{ik}\right) = p_i + \hat{h}_{ik} q_i + \frac{\hat{h}_{ik}^2}{2} u_i + \frac{\hat{h}_{ik}^3}{6} v_i, \tag{3.15}$$

$$\mathbf{x}^i = \left(p_i, q_i, u_i, v_i, s_{i1}, \cdots, s_{iM_i}\right),$$

and

$$\mathbf{x}^n = \left(p_n, q_n\right).$$

The primal variables are represented by

$$\mathbf{x} = \left(\mathbf{x}^0, \mathbf{x}^1, \ldots, \mathbf{x}^{n-1}, \mathbf{x}^n\right) \subset \mathfrak{C} = R^{M+4n+2}, \tag{3.16}$$

which are required to satisfy the conditions (3.10), (3.11), and (3.15). Since these constraints are in linear form, they can be expressed as

$$\mathcal{A}\,\mathbf{x} = 0,$$

where $\mathcal{A}$ is an $(2n + M) \times (4n + M + 2)$ matrix formed by the coefficients in (3.10), (3.11), and (3.15), i.e., $\mathcal{A} =$

$$\left[\begin{array}{cccccccccccccccccccccccc}
1 & h_0 & \frac{h_0^2}{2} & \frac{h_0^3}{6} & & & & -1 \\
 & 1 & h_0 & \frac{h_0^2}{2} & & & & & -1 \\
1 & \hat{h}_{01} & \frac{\hat{h}_{01}^2}{2} & \frac{\hat{h}_{01}^3}{6} & -1 \\
\vdots & & & \vdots & & \ddots \\
1 & \hat{h}_{0M_0} & \frac{\hat{h}_{0M_0}^2}{2} & \frac{\hat{h}_{0M_0}^3}{6} & & & -1 \\
 & & & & & & & 1 & h_1 & \frac{h_1^2}{2} & \frac{h_1^3}{6} & -1 \\
 & & & & & & & & 1 & h_1 & \frac{h_1^2}{2} & & -1 \\
 & & & & & & & & & & & & & \cdots \\
 & & & & & & & & & & & & & & \cdots & 1 & h_{n-1} & \frac{h_{n-1}^2}{2} & \frac{h_{n-1}^3}{6} & & & & -1 \\
 & & & & & & & & & & & & & & \cdots & & 1 & h_{n-1} & \frac{h_{n-1}^2}{2} & & & & & -1 \\
 & & & & & & & & & & & & & & & 1 & \hat{h}_{(n-1)1} & \frac{\hat{h}_{(n-1)1}^2}{2} & \frac{\hat{h}_{(n-1)1}^3}{6} & -1 \\
 & & & & & & & & & & & & & & & \vdots & & & \vdots & & \ddots \\
 & & & & & & & & & & & & & & & 1 & \hat{h}_{(n-1)M_{n-1}} & \frac{\hat{h}_{(n-1)M_{n-1}}^2}{2} & \frac{\hat{h}_{(n-1)M_{n-1}}^3}{6} & & & -1
\end{array}\right] \tag{3.17}$$

We denote the null space of matrix $\mathcal{A}$ by $\mathfrak{X}$. The primal objective function now becomes

$$\mathfrak{g}\left(\mathbf{x}\right) = \sum_{i=0}^{n-1} \mathfrak{g}_i\left(\mathbf{x}^i\right), \tag{3.18}$$

where

$$\mathfrak{g}_i\left(\mathbf{x}^i\right) = \lambda \sum_{k=1}^{M_i} \hat{w}_{ik}\ \left|s_{ik} - \hat{z}_{ik}\right| + (1-\lambda)\ f_i\left(u_i, v_i\right), \tag{3.19}$$

and

$$f_i\left(u_i, v_i\right) = \begin{cases} \frac{h_i}{2}\left[\left|\frac{h_i}{2}v_i\right| + \frac{u_i + \frac{h_i}{2}v_i}{\frac{h_i}{2}v_i}^{2}\right], & \text{if} \quad \left|\frac{h_i}{2}v_i\right| > \left|u_i + \frac{h_i}{2}v_i\right|, \\ h_i\left|u_i + \frac{h_i}{2}v_i\right|, & \text{if} \quad \left|\frac{h_i}{2}v_i\right| \leq \left|u_i + \frac{h_i}{2}v_i\right|. \end{cases} \tag{3.20}$$

Consequently, the primal problem becomes

$$\text{(Primal)} \qquad \begin{cases} \min \mathfrak{g}\left(\mathbf{x}\right) \\ \text{s.t. } \mathbf{x} \in \mathfrak{C} \cap \mathfrak{X} \end{cases}. \tag{3.21}$$

Notice that in this primal problem, there are $4n + M + 2$ unknown variables and $2n + M$ constraints. Consequently, there are $2n + 2$ degrees of freedom.

3.2.2 Dual problem

Following the general theory of [22] and a specific implementation of this theory in [2], we can derive a geometric dual problem for univariate cubic L_1 smoothing splines in the following manner.

Let us denote

$$\mathbf{y}^i = \left(\varphi_1^i, \varphi_2^i, \varphi_3^i, \varphi_4^i, \vartheta_{i1}, \cdots, \vartheta_{iM_i}\right) \quad \text{for} \quad i = 0, \ldots, n-1,$$

$$\mathbf{y}^n = \left(\varphi_1^n, \varphi_2^n\right).$$

The dual variables are represented by

$$\mathbf{y} = \left(\mathbf{y}^0, \mathbf{y}^1, \ldots, \mathbf{y}^{n-1}, \mathbf{y}^n\right). \tag{3.22}$$

The conjugate transform of $\mathfrak{g}\left(\mathbf{x}\right) : \mathfrak{C}$ is denoted by $\mathfrak{h}\left(\mathbf{y}\right) : \mathfrak{D}$. Since the primal function $\mathfrak{g}\left(\mathbf{x}\right)$ is in a separable form, we can derive the conjugate transform of each components independently. By the definition of conjugate transform [24], we have

$$\begin{aligned} \mathfrak{h}_i\left(\mathbf{y}^i\right) \quad &= \quad \sup_{\mathbf{x}^i \in R^{4+M_i}} \{\sum_{k=1}^{M_i}\left[\vartheta_{ik}s_{ik} - \lambda\hat{w}_{ik}\ \left|s_{ik} - \hat{z}_{ik}\right|\right] + \varphi_1^i p_i + \varphi_2^i q_i + \left[\varphi_3^i u_i + \varphi_4^i v_i - (1-\lambda)\, f_i\left(u_i, v_i\right)\right]\} \\ &= \quad \sum_{k=1}^{M_i} \sup_{s_{ik} \in R^1}\left\{\vartheta_{ik}s_{ik} - \lambda\hat{w}_{ik}\ \left|s_{ik} - \hat{z}_{ik}\right|\right\} + \sup_{p_i \in R^1}\left\{\varphi_1^i p_i\right\} + \sup_{q_i \in R^1}\left\{\varphi_2^i q_i\right\} \\ &\quad\quad + \sup_{(u_i, v_i) \in R^2}\left\{\varphi_3^i u_i + \varphi_4^i v_i - (1-\lambda)\ f_i\left(u_i, v_i\right)\right\} \end{aligned}$$

To keep the supremum finite, both φ_1^i and φ_2^i can take on only the value zero. The conjugate transform of $\lambda \hat{w}_{ik} \left| s_{ik} - \hat{z}_{ik} \right| : \{ s_{ik} \in R^1 \}$ is

$$\hat{z}_{ik} \vartheta_{ik} : \{ \left| \vartheta_{ik} \right| \leq \lambda \hat{w}_{ik} \} \tag{3.23}$$

From [2] the conjugate transform of $f_i \left(u_i, v_i \right) : \{ \left(u_i, v_i \right) \in R^2 \}$ is

$$0 : \{ -\frac{h_i^2}{2} + \frac{1}{4} \left(\varphi_3^i + h_i \right)^2 \leq \varphi_4^i \leq \frac{h_i^2}{2} - \frac{1}{4} \left(\varphi_3^i - h_i \right)^2 \} \tag{3.24}$$

or, equivalently,

$$0 : \{ \left| 4\varphi_4^i - 2h_i \varphi_3^i \right| \leq \left(h_i \right)^2 - \left(\varphi_3^i \right)^2 , \quad i = 0, \ldots, n-1 \}. \tag{3.25}$$

Notice that $\left(1 - \lambda \right) f_i \left(u_i, v_i \right) = f_i \left(\left(1 - \lambda \right) u_i, \left(1 - \lambda \right) v_i \right)$. Making the transformations $u_i' = \left(1 - \lambda \right) u_i$ and $v_i' = \left(1 - \lambda \right) v_i$, we have

$$\begin{aligned} &\sup_{\left(u_i, v_i \right) \in R^2} \left\{ \varphi_3^i u_i + \varphi_4^i v_i - \left(1 - \lambda \right) f_i \left(u_i, v_i \right) \right\} \\ &= \sup_{\left(u_i', v_i' \right) \in R^2} \left\{ \frac{\varphi_3^i}{1 - \lambda} u_i' + \frac{\varphi_4^i}{1 - \lambda} v_i' - f_i \left(u_i', v_i' \right) \right\} = 0 \end{aligned}$$

with

$$\left(1 - \lambda \right) \left| 4\varphi_4^i - 2h_i \varphi_3^i \right| \leq \left(1 - \lambda \right)^2 h_i^2 - \left(\varphi_3^i \right)^2 , \quad i = 0, \ldots, n-1. \tag{3.26}$$

Therefore,

$$\mathfrak{h} \left(\mathbf{y} \right) = \sum_{i=0}^{n-1} \sum_{k=1}^{M_i} \hat{z}_{ik} \vartheta_{ik}, \tag{3.27}$$

with a domain

$$\mathfrak{D} = \prod_{i=0}^{n} \mathfrak{D}_i \subset R^{4n+M+2}, \tag{3.28}$$

where, for $i = 0, \ldots, n-1$,

$$\mathfrak{D}_i = \{0\} \times \{0\} \times \Theta_i \times \prod_{k=1}^{M_i} I_{ik}, \tag{3.29}$$

with

$$\Theta_i = \left\{ \left(\varphi_3^i, \varphi_4^i \right) \in R^2 \,\middle|\, \left(1 - \lambda \right) \left| 4\varphi_4^i - 2h_i \varphi_3^i \right| \leq \left(1 - \lambda \right)^2 h_i^2 - \left(\varphi_3^i \right)^2 \right\}, \tag{3.30}$$

$$I_{ik} = \left\{ \vartheta_{ik} \in R^1 \mid \left| \vartheta_{ik} \right| \leq \lambda \hat{w}_{ik} \right\}, \tag{3.31}$$

and, for $i = n$,

$$\mathfrak{D}_n = \{0\} \times \{0\}. \tag{3.32}$$

Therefore, the dual geometric programming problem becomes

$$\text{(Dual)} \qquad \begin{cases} \min \mathfrak{h}\left(\mathbf{y}\right) \\ \\ \text{s.t.} \;\; \mathbf{y} \in \mathfrak{D} \cap \mathfrak{Y} \end{cases}, \tag{3.33}$$

where $\mathfrak{Y}$, the dual cone of $\mathfrak{X}$, is the row space of the matrix $\mathcal{A}$ as defined in (3.17).

3.2.3 A simplified dual program

Let the row vectors of $\mathcal{A}$ be $\boldsymbol{\xi}_i$, $\boldsymbol{\eta}_i$, and $\boldsymbol{\zeta}_{ik}$ for $i = 0, \cdots, n-1$ and $k = 1, \cdots, M_i$, i.e.,

$$\mathcal{A}^T = \left[\begin{array}{ccccccccccc} {}^T_0 & {}^T_0 & {}^T_{01} & \cdots & {}^T_{0M_0} & \cdots & {}^T_{n-1} & {}^T_{n-1} & {}^T_{(n-1)1} & \cdots & {}^T_{(n-1)M_{n-1}} \end{array} \right]. \tag{3.34}$$

For each $\mathbf{y} \in \mathfrak{Y}$, since $\mathfrak{Y}$ is the row space of $\mathcal{A}$, $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \sum_{i=0}^{n-1} \left(\alpha_i \boldsymbol{\xi}_i + \beta_i \boldsymbol{\eta}_i + \sum_{k=1}^{M_i} \gamma_{ik} \boldsymbol{\zeta}_{ik} \right), \tag{3.35}$$

with $\alpha_i, \beta_i, \gamma_{ik} \in R$. From (3.35) we easily have

$$\gamma_{ik} = -\vartheta_{ik}. \tag{3.36}$$

Hence (3.35) can be written explicitly as

$$\begin{aligned} \varphi_1^0 &= \alpha_0 - \sum_{k=1}^{M_0} \vartheta_{0k}, \\ \varphi_1^i &= -\alpha_{i-1} + \alpha_i - \sum_{k=1}^{M_i} \vartheta_{ik}, \;\; i = 1, \ldots, n-1, \\ \varphi_1^n &= -\alpha_{n-1}, \end{aligned} \tag{3.37}$$

$$\begin{aligned} \varphi_2^0 &= h_0 \alpha_0 + \beta_0 - \sum_{k=1}^{M_0} \hat{h}_{0k} \vartheta_{0k}, \\ \varphi_2^i &= -\beta_{i-1} + h_i \alpha_i + \beta_i - \sum_{k=1}^{M_i} \hat{h}_{ik} \vartheta_{ik}, \; i = 1, \ldots, n-1, \\ \varphi_2^n &= -\beta_{n-1}, \end{aligned} \tag{3.38}$$

$$\varphi_3^i = \frac{h_i^2}{2} \alpha_i + h_i \beta_i - \sum_{k=1}^{M_i} \frac{\hat{h}_{ik}^2}{2} \vartheta_{ik}, \;\; i = 0, \ldots, n-1, \tag{3.39}$$

$$\varphi_4^i = \frac{h_i^3}{6} \alpha_i + \frac{h_i^2}{2} \beta_i - \sum_{k=1}^{M_i} \frac{\hat{h}_{ik}^3}{6} \vartheta_{ik}, \;\; i = 0, \ldots, n-1. \tag{3.40}$$

Recalling (3.29) and (3.32), we have $\varphi_1^i = 0$ and $\varphi_2^i = 0$ for $i = 0, \cdots, n$. Hence, from (3.37) and (3.38), α_i and β_i can be expressed as linear functions of the ϑ_{ik}, i.e.,

$$\begin{aligned}
&\alpha_i = \sum_{j=0}^{i} \sum_{k=1}^{M_j} \vartheta_{jk}, \quad i = 0, \ldots, n-1, \\
&\beta_i = -\sum_{j=0}^{i} \sum_{k=1}^{M_j} (x_{i+1} - \hat{x}_{jk})\, \vartheta_{jk}, \quad i = 0, \ldots, n-1, \\
&\alpha_{n-1} = 0, \\
&\beta_{n-1} = 0,
\end{aligned} \tag{3.41}$$

and, for $i = 0, \ldots, n-1$,

$$\begin{aligned}
\varphi_3^i &= -\tfrac{1}{2} \sum_{j=0}^{i-1} \sum_{k=1}^{M_j} h_i \left(x_i + x_{i+1} - 2\hat{x}_{jk}\right) \vartheta_{jk} - \tfrac{1}{2} \sum_{k=1}^{M_i} \left(x_{i+1} - \hat{x}_{ik}\right)^2 \vartheta_{ik} \\
\varphi_4^i &= -\tfrac{1}{6} \sum_{j=0}^{i-1} \sum_{k=1}^{M_j} h_i^2 \left(x_i + 2x_{i+1} - 3\hat{x}_{jk}\right) \vartheta_{jk} - \tfrac{1}{6} \sum_{k=1}^{M_i} \left(x_{i+1} - \hat{x}_{ik}\right)^2 \left(\hat{x}_{ik} + 2x_{i+1} - 3x_i\right) \vartheta_{ik}
\end{aligned} \tag{3.42}$$

Moreover,

$$\left|4\varphi_4^i - 2h_i\varphi_3^i\right| = \left| -\frac{h_i^3}{3} \sum_{j=0}^{i-1} \sum_{k=1}^{M_j} \vartheta_{jk} - \frac{1}{3} \sum_{k=1}^{M_i} \left(x_{i+1} - \hat{x}_{ik}\right)^2 \left(x_{i+1} + 2\hat{x}_{ik} - 3x_i\right) \vartheta_{ik} \right|,$$

for $i = 0, \ldots, n-1$.

Therefore, the geometric dual problem (3.33) can be simplified as

$$\left\{ \begin{array}{ll}
\min\limits_{\vartheta} & \sum\limits_{i=0}^{n-1} \sum\limits_{k=1}^{M_i} \hat{z}_{ik} \vartheta_{ik} \\
\text{s.t.} & |\vartheta_{ik}| \le \lambda \hat{w}_{ik},, \quad i = 0, \ldots, n-1, \qquad k = 1, \cdots, M_i, \\
& \quad (1-\lambda) \left| -\frac{h_i^3}{3} \sum\limits_{j=0}^{i-1} \sum\limits_{k=1}^{M_j} \vartheta_{jk} - \frac{1}{3} \sum\limits_{k=1}^{M_i} \left(x_{i+1} - \hat{x}_{ik}\right)^2 \left(x_{i+1} + 2\hat{x}_{ik} - 3x_i\right) \vartheta_{ik} \right| \\
& \quad \le (1-\lambda)^2 h_i^2 - \left(-\frac{1}{2} \sum\limits_{j=0}^{i-1} \sum\limits_{k=1}^{M_j} h_i \left(x_i + x_{i+1} - 2\hat{x}_{jk}\right) \vartheta_{jk} - \frac{1}{2} \sum\limits_{k=1}^{M_i} \left(x_{i+1} - \hat{x}_{ik}\right)^2 \vartheta_{ik} \right)^2, \\
& \qquad i = 0, \ldots, n-1, \\
& \sum\limits_{j=0}^{n-1} \sum\limits_{k=1}^{M_j} \vartheta_{jk} = 0, \\
& \sum\limits_{j=0}^{n-1} \sum\limits_{k=1}^{M_j} \left(x_n - \hat{x}_{jk}\right) \vartheta_{jk} = 0.
\end{array} \right. \tag{3.43}$$

Notice that this problem has a linear objective function with M variables ϑ_{ik} over a convex feasible domain defined by M box constraints, $2n$ quadratic constraints, and 2 linear constraints. The variables are sequentially related through those quadratic constraints.

3.2.4 Dual-to-primal transformation

After a dual optimal solution $\mathbf{y}^*$ is obtained by solving problem (3.43), a corresponding primal optimal solution can be obtained by the following optimality conditions:

$$\begin{aligned}
&\text{(I)} && \mathbf{x}^* \in \mathfrak{C} \cap \mathfrak{X},\ \mathbf{y}^* \in \mathfrak{D} \cap \mathfrak{Y};\\
&\text{(II)} && \langle \mathbf{x}^*, \mathbf{y}^* \rangle = 0;\\
&\text{(III)} && \mathbf{x}^* \in \partial \mathfrak{h}\left(\mathbf{y}^*\right).
\end{aligned}$$

Since $\mathfrak{X}$ and $\mathfrak{Y}$ are dual cones, the primal optimal solution $\mathbf{x}^*$ is a vector such that

$$\mathbf{x}^* \in \mathfrak{C} \cap \mathfrak{X} \cap \partial \mathfrak{h}\left(\mathbf{y}^*\right). \tag{3.44}$$

From (3.12), (3.13), and (3.15), a feasible primal variable $\mathbf{x} \in \mathfrak{C} \cap \mathfrak{X}$ can be expressed as linear functions of free variables p_i and q_i as follows:

$$u_i = -\frac{6}{h_i^2} p_i - \frac{4}{h_i} q_i + \frac{6}{h_i^2} p_{i+1} - \frac{2}{h_i} q_{i+1}, \tag{3.45}$$

$$v_i = \frac{12}{h_i^3} p_i + \frac{6}{h_i^2} q_i - \frac{12}{h_i^3} p_{i+1} + \frac{6}{h_i^2} q_{i+1}, \tag{3.46}$$

$$s_{ik} = \left. 1 - \frac{3\hat{h}_{ik}^2}{h_i^2} + \frac{2\hat{h}_{ik}^3}{h_i^3} \right) p_i + \left. \hat{h}_{ik} - \frac{2\hat{h}_{ik}^2}{h_i} + \frac{\hat{h}_{ik}^3}{h_i^2} \right) q_i + \left. \frac{3\hat{h}_{ik}^2}{h_i^2} - \frac{2\hat{h}_{ik}^3}{h_i^3} \right) p_{i+1} + \left. -\frac{\hat{h}_{ik}^2}{h_i} + \frac{\hat{h}_{ik}^3}{h_i^2} \right) q_{i+1}. \tag{3.47}$$

Moreover, for any $\bar{\mathbf{y}} \in \mathfrak{D}$, $\mathbf{x} \in \partial \mathfrak{h}\left(\bar{\mathbf{y}}\right)$ can be calculated componentwise by the definition of the subgradient, i.e.,

$$\langle \mathbf{x}, \mathbf{y} - \bar{\mathbf{y}} \rangle \leq \mathfrak{h}\left(\mathbf{y}\right) - \mathfrak{h}\left(\bar{\mathbf{y}}\right),\ \forall \mathbf{y} \in \mathfrak{D}. \tag{3.48}$$

Since φ_1^i and φ_2^i can only be zero, p_i and q_i are free to take on any real values. Assume $\mathbf{y}$ is a vector whose elements are the same as the given vector $\bar{\mathbf{y}}$, except for the component $\vartheta_{ik} = \bar{\vartheta}_{ik} + \delta$. Then

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} - \bar{\mathbf{y}} \rangle &= s_{ik}\delta,\\
\mathfrak{h}\left(\mathbf{y}\right) - \mathfrak{h}\left(\bar{\mathbf{y}}\right) &= \hat{z}_{ik}\delta.
\end{aligned}$$

If $\bar{\vartheta}_{ik}$ is located inside the open interval $(-\lambda\hat{w}_{ik}, \lambda\hat{w}_{ik})$, then δ can be either a positive or a negative real number. Hence, we have $s_{ik} = \hat{z}_{ik}$. If $\bar{\vartheta}_{ik}$ is the left end point $-\lambda\hat{w}_{ik}$, then δ can only be a positive number. Thus, s_{ik} must satisfy $s_{ik} \leq \hat{z}_{ik}$. Similarly, if $\bar{\vartheta}_{ik}$ is the right end point $\lambda\hat{w}_{ik}$, we obtain $s_{ik} \geq \hat{z}_{ik}$. All these three cases can be written as

$$s_{ik} = \hat{z}_{ik} + \hat{c}_{ik} t_{ik}, \tag{3.49}$$

where $t_{ik} \geq 0$ and $\hat{c}_{ik}$ is a constant determined by $\hat{c}_{ik} = sign\left(\vartheta_{ik}\right)\left[1 - I_{\left(-\lambda \hat{w}_{ik}, \lambda \hat{w}_{ik}\right)}\left(\vartheta_{ik}\right)\right]$.

By the same logic, we can verify that (u_i, v_i) lies in the normal cone of Θ_i at the point $\left(\bar{\varphi}_3^i, \bar{\varphi}_4^i\right)$. Therefore, (u_i, v_i) can be expressed as

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} = \begin{pmatrix} c_1^i \\ c_2^i \end{pmatrix} r_i + \begin{pmatrix} c_3^i \\ c_4^i \end{pmatrix} s_i \quad \text{for some} \quad r_i, s_i \geq 0, \tag{3.50}$$

Recall that $\left(\bar{\varphi}_3^i, \bar{\varphi}_4^i\right)$ is defined on the set

$$\Theta_i = \left\{ \left(\varphi_3^i, \varphi_4^i\right) \in R^2 \,\middle|\, (1-\lambda)\left|4\varphi_4^i - 2h_i\varphi_3^i\right| \leq (1-\lambda)^2 h_i^2 - \left(\varphi_3^i\right)^2 \right\},$$

which is a convex set bounded by two quadratic curves:

$$C_1^i: \ \ \bar{\varphi}_4^i = -\frac{1-\lambda}{2} h_i^2 + \frac{1}{4(1-\lambda)}\left[\bar{\varphi}_3^i + (1-\lambda) h_i\right]^2$$

and

$$C_2^i: \ \ \bar{\varphi}_4^i = \frac{1-\lambda}{2} h_i^2 - \frac{1}{4(1-\lambda)}\left[\bar{\varphi}_3^i - (1-\lambda) h_i\right]^2.$$

Therefore, $\left(c_1^i, c_2^i\right) = \left(\frac{\bar{\varphi}_3^i}{2(1-\lambda)} + \frac{h_i}{2}, -1\right)$, if $\left(\bar{\varphi}_3^i, \bar{\varphi}_4^i\right)$ is on the curve C_1^i. Otherwise $\left(c_1^i, c_2^i\right) = (0,0)$. And $\left(c_3^i, c_4^i\right) = \left(\frac{\bar{\varphi}_3^i}{2(1-\lambda)} - \frac{h_i}{2}, 1\right)$, if $\left(\bar{\varphi}_3^i, \bar{\varphi}_4^i\right)$ is on the curve C_2^i. Otherwise $\left(c_3^i, c_4^i\right) = (0,0)$.

Putting all the above expressions together, a primal optimal solution $\mathbf{x}^*$ lying in the set $\mathfrak{C} \cap \mathfrak{X} \cap \partial\mathfrak{h}\left(\mathbf{y}^*\right)$ can be obtained by solving the following problem:

$$\left\{ \begin{array}{lll} \min\limits_{p_i, q_i, r_i, s_i, t_{ik}} & \sum_{i=0}^n |q_i| & \\ \text{s.t.} & -\frac{6}{h_i^2} p_i - \frac{4}{h_i} q_i + \frac{6}{h_i^2} p_{i+1} - \frac{2}{h_i} q_{i+1} = c_1^i r_i + c_3^i s_i, & \\ & \frac{12}{h_i^3} p_i + \frac{6}{h_i^2} q_i - \frac{12}{h_i^3} p_{i+1} + \frac{6}{h_i^2} q_{i+1} = c_2^i r_i + c_4^i s_i, & \\ & \left(1 - \frac{3\hat{h}_{ik}^2}{h_i^2} + \frac{2\hat{h}_{ik}^3}{h_i^3}\right) p_i + \left(\hat{h}_{ik} - \frac{2\hat{h}_{ik}^2}{h_i} + \frac{\hat{h}_{ik}^3}{h_i^2}\right) q_i + \left(\frac{3\hat{h}_{ik}^2}{h_i^2} - \frac{2\hat{h}_{ik}^3}{h_i^3}\right) p_{i+1} & \\ & \quad + \left(-\frac{\hat{h}_{ik}^2}{h_i} + \frac{\hat{h}_{ik}^3}{h_i^2}\right) q_{i+1} = \hat{z}_{ik} + \hat{c}_{ik} t_{ik}, & \\ & r_i, s_i, t_{ik} \geq 0, & \\ & p_i, q_i \text{ unrestricted.} & \end{array} \right. \tag{3.51}$$

where the objective function is the regularization term mentioned in [2]. One could, of course, use other objective functions such as $\sum_{i=0}^n q_i^2$.

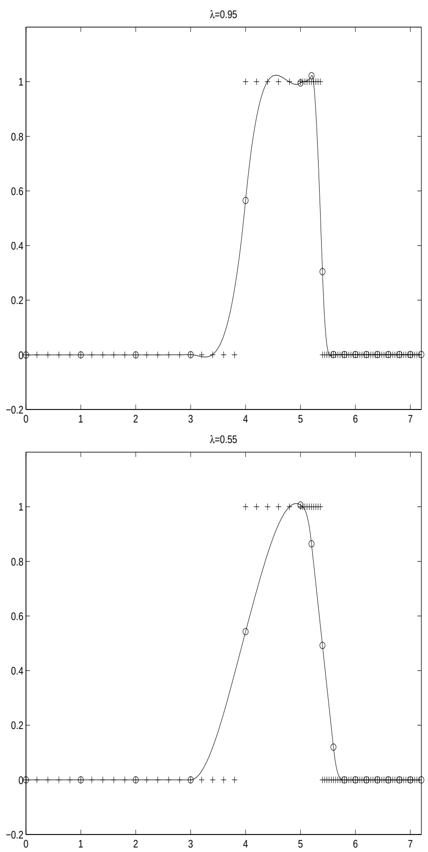

Figure 3.1: Cubic L_1 smoothing splines for a plateau.

3.3 Numerical experiments

In this section, we use two examples of [15] to illustrate the results obtained by the geometric programming approach. Figure 1 is for data with a plateau while Figure 2 is for data with flat portions, linear ramps, and an outlier. In both figures, the data points to be approximated are indicated by the symbol "+" and the spline knots are indicated by the symbol "∘". The weights $\hat{w}_m$ of the data points are all set equal to 1. For each example, four values of the balance parameter λ are used. The solver used for the simplified dual problem was basic MINOS 5.5 without any modification.

For both examples, the results obtained by the geometric programming approach are essentially the same as the results obtained by the discretization approach of [8]. The real advantage of the geometric programming approach lies in the theoretical analysis to be presented in the next section.

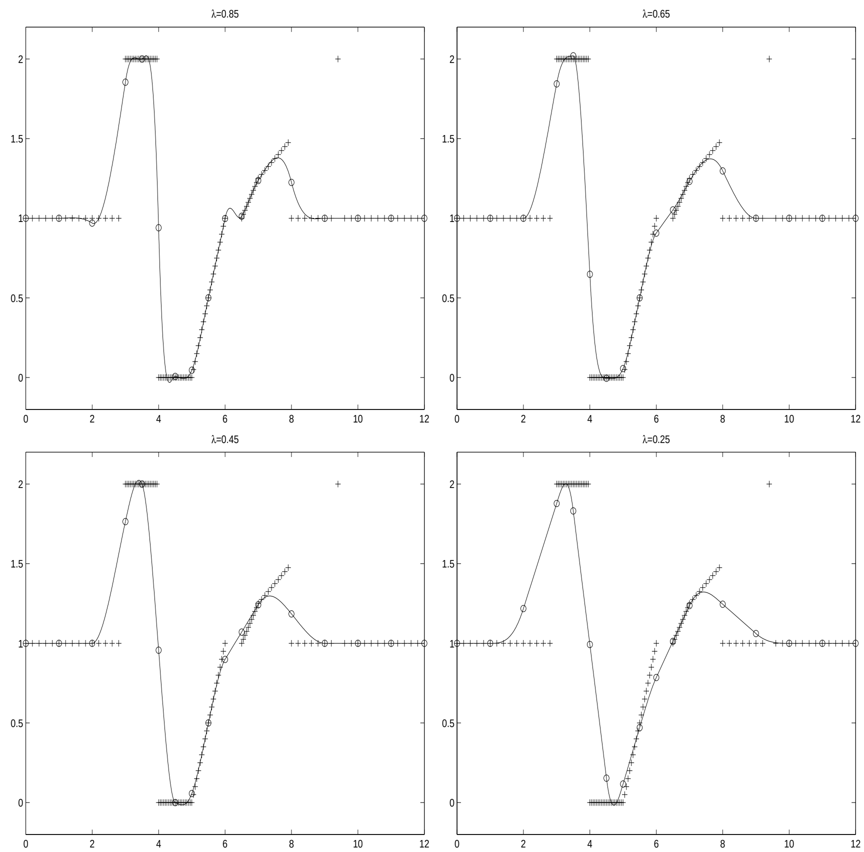

Figure 3.2: Cubic L_1 smoothing splines for data with flat portions, linear ramps and an outlier at $(9.4, 2)$.

3.4 Sensitivity analysis for data with uncertainty

So far we have assumed that every data value $\hat{z}$ is given in a precise manner. In reality, for many applications, many data points may come with uncertainty represented by an interval. In this case, some theoretical analysis is needed.

Let us assume that, at a given location $\hat{x}_k$, the data value $\hat{z}_k$ is subject to measurement error. Then $\hat{z}_k$ is no longer a precise constant, but may assume any value in an interval $[z_l, z_u]$. For simplicity, let us skip the subscript and denote this data $(\hat{x}, z)$ with $z \in [z_l, z_u]$.

For each value of z, we are able to find one corresponding cubic L_1 smoothing spline as described in Section 2. In this situation, the objective function of the unconstrained optimization problem (3.14) becomes $E(\mathbf{p}, \mathbf{q}; z)$, where $\mathbf{p}, \mathbf{q} \in R^{n+1}$ and $z \in [z_l, z_u]$. Consequently, its objective function value becomes a function in term of z, i.e.,

$$\widetilde{E}(z) = \min_{\mathbf{p},\mathbf{q}\in R^{n+1}} E(\mathbf{p},\mathbf{q};z). \tag{3.52}$$

The data value z is thus a parameter that defines a family of smoothing splines. Among this smoothing spline family, we are particularly interested in finding the smoothing spline that has the minimum objective function value. That is, we face the following optimization problem

$$\min_{z\in[z_l,z_u]} \widetilde{E}(z) = \min_{z\in[z_l,z_u]} \min_{\mathbf{p},\mathbf{q}\in R^{n+1}} E(\mathbf{p},\mathbf{q};z)\ . \tag{3.53}$$

In function $E(\mathbf{p},\mathbf{q};z)$, z appears only in the term $\lambda\hat{w}\ |S(\hat{x}) - z|$, and inside the absolute value symbol is a linear function of $\mathbf{p},\mathbf{q}$ and z. Hence this term is in fact a convex function of $\mathbf{p},\mathbf{q}$ and z. We have shown in Section 2 that, without this term, $E(\mathbf{p},\mathbf{q};z)$ is a convex function of $\mathbf{p}$ and $\mathbf{q}$. Therefore, $E(\mathbf{p},\mathbf{q};z)$ is also a convex function [24].

Now we show that $\widetilde{E}(z)$ is convex. For any z_1, $z_2 \in [z_l, z_u]$ and $\alpha \in [0,1]$, let $(\mathbf{p}_1,\mathbf{q}_1)$, $(\mathbf{p}_2,\mathbf{q}_2)$, and $(\mathbf{p}^*,\mathbf{q}^*)$ be minimizers of $E(\mathbf{p},\mathbf{q};z)$ with respect to $z = z_1$, $z = z_2$, and $z = \alpha z_1 + (1-\alpha) z_2$. We have

$$\begin{aligned}
&\widetilde{E}(\alpha z_1 + (1-\alpha) z_2) \\
&= E(\mathbf{p}^*,\mathbf{q}^*;\alpha z_1 + (1-\alpha) z_2) \\
&\le E(\alpha\mathbf{p}_1 + (1-\alpha)\mathbf{p}_2, \alpha\mathbf{q}_1 + (1-\alpha)\mathbf{q}_2; \alpha z_1 + (1-\alpha) z_2) \\
&\le \alpha E(\mathbf{p}_1,\mathbf{q}_1;z_1) + (1-\alpha) E(\mathbf{p}_2,\mathbf{q}_2;z_2) \\
&= \alpha\widetilde{E}(z_1) + (1-\alpha)\widetilde{E}(z_2)
\end{aligned}$$

Consequently, we have the following theorem:

Theorem 3.1 *$\widetilde{E}(z)$ is a convex function over $[z_l, z_u]$.*

For the geometric programming model of the L_1 smoothing splines, let us use $\mathfrak{g}(\mathbf{x};z)$ and $\mathfrak{h}(\mathbf{y};z)$ to denote the primal and dual objective functions, respectively, with z being an additional variable. Also define that

$$\widetilde{\mathfrak{g}}(z) = \min_{\mathbf{x} \in \mathfrak{C} \cap \mathfrak{X}} \mathfrak{g}(\mathbf{x}; z),$$

$$\widetilde{\mathfrak{h}}(z) = \min_{\mathbf{y} \in \mathfrak{D} \cap \mathfrak{Y}} \mathfrak{h}(\mathbf{y}; z).$$

Following the logic previously used, it is not difficult to show that $\widetilde{\mathfrak{g}}(z)$ is convex. By the optimality theorem [21] of geometric programming, we know that for any $z \in [z_l, z_u]$,

$$\widetilde{\mathfrak{g}}(z) + \widetilde{\mathfrak{h}}(z) = 0. \tag{3.54}$$

Then it is straightforward to see that $\widetilde{\mathfrak{h}}(z)$ is concave.

Corollary 3.2 $\widetilde{\mathfrak{g}}(z)$ *is a convex function on* $[z_l, z_u]$*,* $\widetilde{\mathfrak{h}}(z)$ *is a concave function on* $[z_l, z_u]$*, and minimizing* $\widetilde{\mathfrak{g}}(z)$ *is equivalent to maximizing* $\widetilde{\mathfrak{h}}(z)$*.*

We now change the dual variable from $\mathbf{y}$ to $\boldsymbol{\vartheta}$ as described in Section 3.2.3. In this setting, the feasible domain of $\boldsymbol{\vartheta}$ is a convex set $\Omega_{\boldsymbol{\vartheta}}$ defined in (3.43), which is independent of z. For any $z \in [z_l, z_u]$, the function $\widetilde{\mathfrak{h}}(z)$ becomes

$$\widetilde{\mathfrak{h}}(z) = \widetilde{\mathbf{z}}^T \widetilde{\boldsymbol{\vartheta}} \tag{3.55}$$

where $\widetilde{\mathbf{z}} = (\hat{z}_1, \ldots, \hat{z}_{k-1}, z, \hat{z}_{k+1}, \ldots, \hat{z}_M)^T$ is a vector of the values of given data, the kth component of which is the value of z, and $\widetilde{\boldsymbol{\vartheta}} = \left(\hat{\vartheta}_1, \ldots, \hat{\vartheta}_{k-1}, \vartheta, \hat{\vartheta}_{k+1}, \ldots, \hat{\vartheta}_M\right)^T$ is one corresponding optimal dual solution, the kth component of which is ϑ. Since every element of $\widetilde{\mathbf{z}}$ is fixed except the kth element z, $\widetilde{\boldsymbol{\vartheta}}$ is determined by the value of z. We denote the kth element of $\widetilde{\boldsymbol{\vartheta}}$ by $\vartheta(z)$. Note that, in general, $\vartheta(z)$ is a set-valued function since there may exist multiple optimal solutions for a given z. In the following discussion, $\vartheta(z)$ may refer to the whole set or any real number in this set. The reader should have no difficulty in determining the meaning from the context.

For any z_1, $z_2 \in [z_l, z_u]$, let $\widetilde{\mathbf{z}}_1, \widetilde{\mathbf{z}}_2$ represent the corresponding vectors of data values. The only difference between $\widetilde{\mathbf{z}}_1$ and $\widetilde{\mathbf{z}}_2$ is in the kth elements. Let the bold characters $\widetilde{\boldsymbol{\vartheta}}_1$ and $\widetilde{\boldsymbol{\vartheta}}_2$ represent the corresponding optimal dual solutions with kth elements $\vartheta(z_1)$ and $\vartheta(z_2)$, respectively. Since, for fixed z_1, the dual optimal solution is $\widetilde{\boldsymbol{\vartheta}}_1$, we know that

$$\widetilde{\mathbf{z}}_1^T \widetilde{\boldsymbol{\vartheta}}_1 \leq \widetilde{\mathbf{z}}_1^T \widetilde{\boldsymbol{\vartheta}}_2. \tag{3.56}$$

Similarly, for z_2,

$$\widetilde{\mathbf{z}}_2^T \widetilde{\boldsymbol{\vartheta}}_2 \leq \widetilde{\mathbf{z}}_2^T \widetilde{\boldsymbol{\vartheta}}_1. \tag{3.57}$$

By inequalities (3.56) and (3.57), we have

$$\begin{aligned} & \widetilde{\mathbf{z}}_1^T \widetilde{\boldsymbol{\vartheta}}_1 = \widetilde{\mathbf{z}}_2^T \widetilde{\boldsymbol{\vartheta}}_1 + (\widetilde{\mathbf{z}}_1 - \widetilde{\mathbf{z}}_2)^T \widetilde{\boldsymbol{\vartheta}}_1 \\ \leq \ & \widetilde{\mathbf{z}}_1^T \widetilde{\boldsymbol{\vartheta}}_2 = \widetilde{\mathbf{z}}_2^T \widetilde{\boldsymbol{\vartheta}}_2 + (\widetilde{\mathbf{z}}_1 - \widetilde{\mathbf{z}}_2)^T \widetilde{\boldsymbol{\vartheta}}_2 \\ \leq \ & \widetilde{\mathbf{z}}_2^T \widetilde{\boldsymbol{\vartheta}}_1 + (\widetilde{\mathbf{z}}_1 - \widetilde{\mathbf{z}}_2)^T \widetilde{\boldsymbol{\vartheta}}_2. \end{aligned}$$

Consequently,

$$(\widetilde{\mathbf{z}}_1 - \widetilde{\mathbf{z}}_2)^T \left(\widetilde{\boldsymbol{\vartheta}}_1 - \widetilde{\boldsymbol{\vartheta}}_2 \right) \leq 0$$

which is equivalent to

$$(z_1 - z_2) \cdot (\vartheta(z_1) - \vartheta(z_2)) \leq 0. \tag{3.58}$$

Hence ϑ is monotonically nonincreasing in z.

Assume that z^* is the maximizer of $\widetilde{\mathfrak{h}}(z)$. The value z^* must exist, since $\mathfrak{h}(\mathbf{y}; z)$ and, hence, $\widetilde{\mathfrak{h}}(z)$ is continuous over the closed interval $[z_l, z_u]$. Denote the optimal dual solution corresponding to the value of z^* by $\widetilde{\boldsymbol{\vartheta}}^*$. In other words,

$$\widetilde{\mathfrak{h}}(z^*) = \max_{z \in [z_l, z_u]} \widetilde{\mathfrak{h}}(z) = \max_{z \in [z_l, z_u]} \min_{\boldsymbol{\vartheta} \in \Omega_{\boldsymbol{\vartheta}}} \widetilde{\mathbf{z}}^T \boldsymbol{\vartheta} = (\widetilde{\mathbf{z}}^*)^T \widetilde{\boldsymbol{\vartheta}}^*. \tag{3.59}$$

Since z^* is the maximizer of $\widetilde{\mathfrak{h}}(z)$, for any $z \in [z_l, z_u]$, we have

$$\widetilde{\mathfrak{h}}(z) = \widetilde{\mathbf{z}}^T \widetilde{\boldsymbol{\vartheta}} \leq (\widetilde{\mathbf{z}}^*)^T \widetilde{\boldsymbol{\vartheta}}^* = \widetilde{\mathfrak{h}}(z^*). \tag{3.60}$$

Furthermore, since $\widetilde{\boldsymbol{\vartheta}}^*$ is the dual optimal solution corresponding to z^*, we have

$$(\widetilde{\mathbf{z}}^*)^T \widetilde{\boldsymbol{\vartheta}}^* \leq (\widetilde{\mathbf{z}}^*)^T \widetilde{\boldsymbol{\vartheta}}. \tag{3.61}$$

Combining (3.60) and (3.61), we have

$$\widetilde{\mathbf{z}}^T \widetilde{\boldsymbol{\vartheta}} \leq (\widetilde{\mathbf{z}}^*)^T \widetilde{\boldsymbol{\vartheta}}^* \leq (\widetilde{\mathbf{z}}^*)^T \widetilde{\boldsymbol{\vartheta}}.$$

Hence,

$$(\widetilde{\mathbf{z}} - \widetilde{\mathbf{z}}^*)^T \widetilde{\boldsymbol{\vartheta}} \leq 0,$$

or equivalently,

$$(z - z^*) \cdot \vartheta(z) \leq 0. \tag{3.62}$$

The above derivation indicates that z^* can be categorized through $\vartheta(z)$ using (3.58) and (3.62). First let us consider the case that z^* is an endpoint of $[z_l, z_u]$. Without loss of generality, let us assume that $z^* = z_u$. From (3.62), $z - z^* \leq 0$ for any $z \in [z_l, z_u]$. Consequently, $\vartheta(z) \geq 0$. At the two end points, we have $\vartheta(z_l) \geq 0$ and $\vartheta(z_u) \geq 0$. On the other hand, if we know that $\vartheta(z_l), \vartheta(z_u) \geq 0$, by the monotonicity property (3.58), we know that $\vartheta(z) \geq 0$ for any $z \in [z_l, z_u]$. Hence, by (3.62), $z \leq z^*$. Therefore, we must have $z^* = z_u$. Therefore, the sufficient and necessary conditions for $z^* = z_u$ are $\vartheta(z_l) \geq 0$ and $\vartheta(z_u) \geq 0$. Similarly, we can show that $z^* = z_l$ if and only if $\vartheta(z_l) \leq 0$ and $\vartheta(z_u) \leq 0$.

Now consider the case that z^* is an interior point of $[z_l, z_u]$. In this case $\vartheta(z_l)$ and $\vartheta(z_u)$ have different signs. If fact, it is not difficult to see from the monotonicity condition that $\vartheta(z_l) \geq 0$ and $\vartheta(z_u) \leq 0$. Let us first assume $\vartheta(z^*) > 0$. Then there exists $z \in [z_l, z_u]$ such that $0 \leq \vartheta(z) < \vartheta(z^*)$. From (3.62), we know that $z - z^* < 0$. However, this implies $(z - z^*) \cdot (\vartheta(z) - \vartheta(z^*)) > 0$, which contradicts (3.58). Therefore, any value of z with $\vartheta(z) > 0$ can not be the maximizer of $\widetilde{\mathfrak{h}}(z)$. By a similar argument, any value of z with $\vartheta(z) < 0$ can not be the maximizer of $\widetilde{\mathfrak{h}}(z)$ either. Hence, z^* is the maximizer of $\widetilde{\mathfrak{h}}(z)$ only if $0 \in \vartheta(z^*)$. On the other hand, since the maximizer always exists in the interval (z_l, z_u), this necessary condition is also a sufficient condition. This characterization is summarized in the following theorem:

Theorem 3.3 *The maximizer z^* of $\widetilde{\mathfrak{h}}(z)$ must exist in $[z_l, z_u]$. In particular, (i) $z^* = z_u$ if and only if $\vartheta(z_l) \geq 0$ and $\vartheta(z_u) \geq 0$; (ii) $z^* = z_l$ if and only if $\vartheta(z_l) \leq 0$ and $\vartheta(z_u) \leq 0$; and (iii) if $\vartheta(z_l)$ and $\vartheta(z_u)$ have different signs, then z^* is the maximizer of $\widetilde{\mathfrak{h}}(z)$ if and only if $0 \in \vartheta(z^*)$.*

The following is an algorithm for finding z^*.

Algorithm 3.1

Given a sufficiently small number $\epsilon > 0$.

Step 1: Solve the dual problem (3.43) with $z = z_l$ for $\vartheta(z_l)$ and $z = z_u$ for $\vartheta(z_u)$.

Step 2: If $\vartheta(z_l) \geq 0$ and $\vartheta(z_u) \geq 0$, then $z^* = z_u$ and stop. If $\vartheta(z_l) \leq 0$ and $\vartheta(z_u) \leq 0$, then $z^* = z_l$ and stop. If $\vartheta(z_l)$ and $\vartheta(z_u)$ have different signs, continue.

Step 3: Pick any point z_0 between z_l and z_u, solve the corresponding dual problem for $\vartheta(z_0)$. If $|\vartheta(z_0)| < \epsilon$, then $z^* = z_0$ and stop. Otherwise, if $\vartheta(z_0) > 0$, then set $z_l = z_0$, otherwise, set $z_u = z_0$.

Step 4: If $z_u - z_l < \epsilon$, then stop. Otherwise, go to step 3.

The efficiency of this algorithm is determined by the method by which the point z_0 in Step 3 is chosen. Some elementary methods include the bisection approach and interpolation approach. The bisection approach takes $z_0 = (z_l + z_u)/2$, while the interpolation approach is to interpolate points $(z_l, \vartheta(z_l))$ and $(z_u, \vartheta(z_u))$ with a linear function $l(z)$, then find the point z_0 such that $l(z_0) = 0$, i.e.,

$$z_0 = z_l \left(1 - \frac{z_u - z_l}{\vartheta_u - \vartheta_l}\right).$$

There may be more efficient methods.

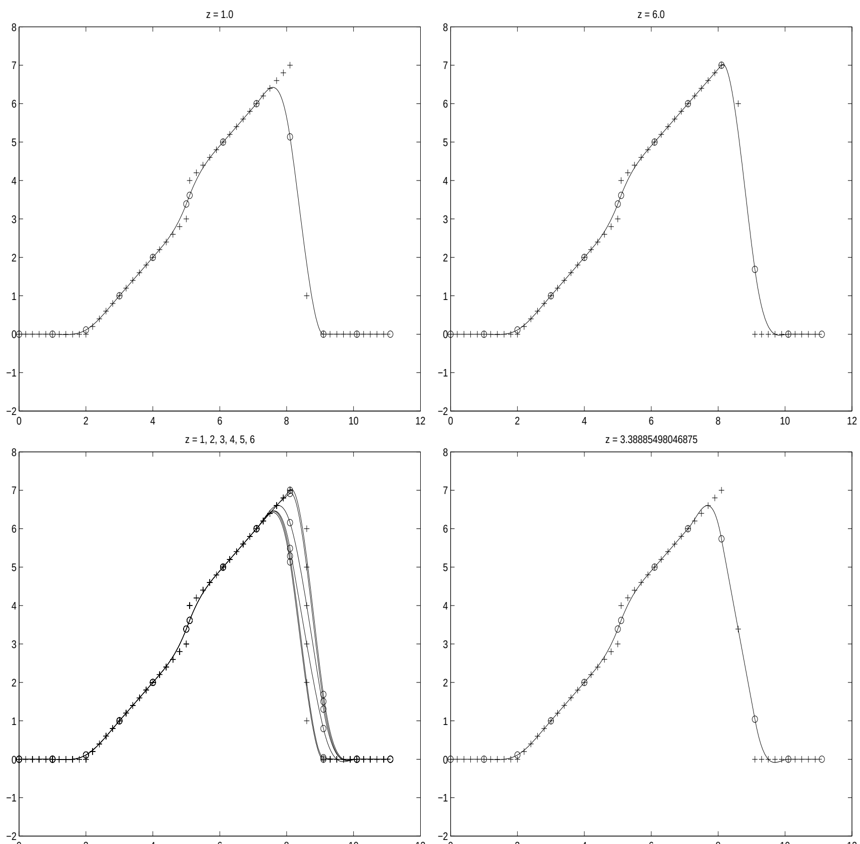

Figure 3.3: Cubic L_1 smoothing splines with $\lambda = 0.75$ for data with linear ramps, vertical glitch, abrupt changes in knot spacing, and various values at $\hat{x} = 8.6$.

We use one example to illustrate the above algorithm. As shown in Figure 3, the data points to be approximated are indicated by the symbol "+" and the spline knots are indicated by the symbol "$\circ$". The weights $\hat{w}_m$ of all data are set to be 1. The balance parameter is fixed at $\lambda = 0.75$. At $\hat{x} = 8.6$, z varies. Numerical results for $z = 1, 2, 3, 4, 5$, and 6 are showed in the following table

z	$\vartheta(z)$	$\widetilde{\mathfrak{h}}(z)$
1.0	0.7	-9.322850347
2.0	0.4600137610105	-8.650202663
3.0	0.4600137607489	-8.190188903
4.0	-0.4600137608493	-8.295175252
5.0	-0.4600137611375	-8.755189014
6.0	-0.7	-9.411460349

The smoothing splines corresponding to $z = 1$ and 6 are shown at the top of Figure 3.3. All 6 smoothing splines are shown at the bottom left of the same figure. From the concavity of $\widetilde{\mathfrak{h}}(z)$, it is clear that the maximizer of $\widetilde{\mathfrak{h}}(z)$ must lie inside the interval $[3, 4]$. Then we use the bisection approach to find that $z^* = 3.38885498046875$. The corresponding smoothing spline is shown in the bottom-right of Figure 3.3.

3.5 Conclusion

We have derived a generalized geometric programming model for the univariate cubic L_1 smoothing spline problem. This formulation provides an attractive theoretical framework for calculating the coefficients of univariate cubic L_1 smoothing splines with or without uncertainty in data. We have also explicitly constructed a geometric dual problem and corresponding optimality conditions. The geometric programming approach provides an exact optimal solution, which does not depend on numerical approximation scheme or human interactions.

Chapter 4

Shape-preserving properties of univariate cubic L_1 splines

In geometric modeling, one of the essential requirements for interpolants is that they "preserve shape well". Shape preservation has often been associated with preserving various properties, such as monotonicity, linearity and convexity/concavity [20]. However, there is no widely accepted definition of shape preservation. It is generally agreed that shape-preserving involves eliminating extraneous non-physical oscillation. These oscillations can be reduced by a variety of techniques, including adjusting the knot positions, introducing additional knots, and adding constraints. However, implementation of these techniques usually requires review of the raw data and human interaction.

Cubic L_1 spline[14, 16] has been shown to be promising for user-input free, shape-preserving interpolating. Empirical experiences support that they are cable of providing C^1-smooth shape-preserving interpolation for arbitrary multi-scale data without additional constraints or user interference. However, this support comes from numerical experiences only, not a theoretic proof. It is not clear if they may fail to preserve shape in some untested circumstances. In fact, Lavery[14] has given examples that a cubic L_1 spline may not preserve traditional shape properties. Even so, many users still feel that the cubic L_1 spline serves their purpose.

In this chapter, we provide a theoretic analysis on the shape-preserving properties of cubic

L_1 splines by exploiting the geometric programming developed in Chapter 2. In particular, we prove that cubic L_1 splines indeed eliminate non-physical oscillations and preserve local linearity and convexity/concavity under mild assumptions for multi-scale data.

4.1 Problem description and notations

4.1.1 Geometric programming formulation

Within geometric programming framework, calculating the coefficients of the univariate cubic L_1 splines is consisted of two steps. The first step is to solve a geometric dual problem to get a dual optimal solution. The second one is to solve the dual-to-primal transformation problem to obtain the primal optimal solution.

By defining

$$b_i = \Delta z_i - \Delta z_{i+1}, i = 0, \ldots, n-2, \tag{4.1}$$

$$\mathbf{B} = (b_0, \ldots, b_{n-2})^T, \tag{4.2}$$

$$f(x) = |x| - \frac{\sqrt{40}-2}{3}, \tag{4.3}$$

and

$$g(x, y) = \frac{1}{3}|x-y| - 1 + \frac{1}{4}(x+y)^2, \tag{4.4}$$

the geometric dual problem is to solve

$$(\mathrm{D}) \quad \begin{cases} \min\limits_{\beta_i} & \sum\limits_{i=0}^{n-2} b_i \beta_i = \langle \mathbf{B}, \beta \rangle \\ \text{s.t.} & f(\beta_0) \leq 0, \\ & g(\beta_{i-1}, \beta_i) \leq 0, \quad i = 1, \ldots, n-2, \\ & f(\beta_{n-2}) \leq 0. \end{cases}$$

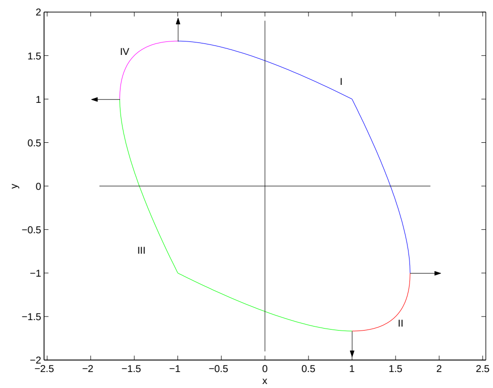

Figure 4.1: Feasible set Ω and four sections of its boundary.

Once the dual optimal solution $\boldsymbol{\beta}^*$ is obtained, the primal optimal solution is given by solving

$$
(\mathrm{T}) \quad \left\{ \begin{array}{lrcl}
\min\limits_{q_i, \lambda_i, \mu_i} & \sum_{i=0}^{n} |q_i| & & \\
\text{s.t.} & q_0 & = & \Delta z_0 - n_1\left(\beta_0^*\right), \\
 & q_1 & = & \Delta z_0 + n_2\left(\beta_0^*\right), \\
 & q_i & = & \Delta z_i - n_1\left(\beta_{i-1}^*, \beta_i^*\right),\ i = 1, \ldots, n-2, \\
 & q_{i+1} & = & \Delta z_i + n_2\left(\beta_{i-1}^*, \beta_i^*\right),\ i = 1, \ldots, n-2, \\
 & q_{n-1} & = & \Delta z_{n-1} - n_1\left(\beta_{n-2}^*\right), \\
 & q_n & = & \Delta z_{n-1} + n_2\left(\beta_{n-2}^*\right), \\
 & \lambda_i, \mu_i & \geq & 0,\ i = 0, \ldots, n-1.
\end{array} \right.
$$

All shape-preserving properties discussed in this chapter are true no matter what is the objective function of the problem (T).

4.1.2 Notations

We begin with introducing some definitions and notations. Let Ω be the feasible set of a pair of consecutive dual variables (β_{i-1}, β_i), $i = 1, \ldots n-2$, i.e.,

$$
\begin{aligned}
\Omega &= \left\{ (x, y)^T \in R^2 \,|\, g(x, y) \leq 0 \right\} \\
&= \left\{ (x, y)^T \in R^2 \,\middle|\, \frac{1}{3}|x - y| \leq 1 - \frac{1}{4}(x + y)^2 \right\}.
\end{aligned}
$$

Let $(n_1(x,y), n_2(x,y))^T$ be the normal vector of Ω at the point (x,y). According to the signs of $n_1(x,y)$ and $n_2(x,y)$, the boundary of Ω is partitioned into four sections. The first section is defined by

$$\text{Section I:} \qquad \left\{ (x,y)^T \in R^2 \middle| -1 \le x \le \frac{5}{3}, -1 \le y \le \frac{5}{3},\ g(x,y) = 0 \right\}.$$

In this way, for any (x,y) in section I, we have $n_1(x,y) \ge 0$ and $n_2(x,y) \ge 0$. The other three sections are defined by

$$\text{Section II:} \qquad \left\{ (x,y)^T \in R^2 \middle| 1 \le x \le \frac{5}{3}, -\frac{5}{3} \le y \le -1,\ g(x,y) = 0 \right\},$$

$$\text{Section III:} \qquad \left\{ (x,y)^T \in R^2 \middle| -\frac{5}{3} \le x \le 1, -\frac{5}{3} \le y \le 1,\ g(x,y) = 0 \right\},$$

and

$$\text{Section IV:} \qquad \left\{ (x,y)^T \in R^2 \middle| -\frac{5}{3} \le x \le -1, 1 \le y \le \frac{3}{5},\ g(x,y) = 0 \right\}.$$

Similarly, if (x,y) is in section II, then $n_1(x,y) \ge 0$ and $n_2(x,y) \le 0$; if (x,y) is in section III, then $n_1(x,y) \le 0$ and $n_2(x,y) \le 0$; and if (x,y) is in section IV, then $n_1(x,y) \le 0$ and $n_2(x,y) \ge 0$.

The feasible set Ω and the four sections of its boundary are illustrated in Figure 4.1.

4.2 Shape-preserving properties

4.2.1 Data dependency

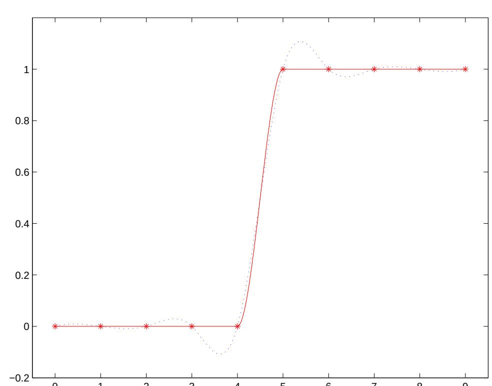

Figure 4.2: Cubic L_1 and L_2 splines of the same data set.

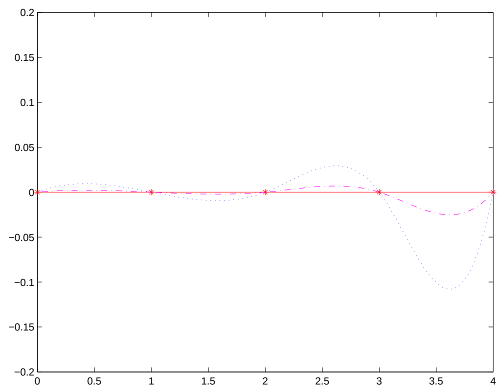

Figure 4.3: Comparison of the data dependency of cubic L_1 and L_2 splines.

For spline interpolation, the data set is given as $\{(x_i, z_i)\}_{i=0}^n$. However, neither x_i nor z_i appears in the geometric programming setting of the cubic L_1 splines. The dual problem (D) and the dual-to-primal transformation (T) are defined by using the divided difference $\{\Delta z_i\}_{i=0}^{n-1}$ only. This unique feature cuts the data dependency on $\{(x_i, z_i)\}_{i=0}^n$.

An example is illustrated in Figure 4.2. The data set $\{(x_i, z_i)\}_{i=0}^n$ is given as $\{(0,0),\ (1,0),$ $(2,0),\ (3,0),\ (4,0),\ (5,1),\ (6,1),\ (7,1),\ (8,1),\ (9,1)\}$. The solid curve is the cubic L_1 spline corresponding to the given data set, while the dotted curve is the corresponding cubic L_2 spline. After the data set is changed to $\{(0,0),\ (1,0),\ (2,0),\ (3,0),\ (4,0),\ (14,10),\ (15,10),\ (16,10),\ (17,10),\ (18,10)\}$, the divided difference $\{\Delta z_i\}$ keeps unchanged. The optimal solution to the L_1 spline problem remains the same, while the L_2 spline changes significantly. Figure 4.3 shows the details of these curves over $[0,4]$. The solid curve represents the cubic L_1 spline, which keeps unchanged, while the dotted and dashed curves of the cubic L_2 splines for two data sets vary significantly.

Note that piecewise linear interpolating function is a perfect multi-scale shape-preserving interpolator. The shape of a piecewise linear interpolation function is completely reflected by $\{\Delta z_i\}_{i=0}^{n-1}$. If we consider the curve as "grid independence", i.e., the axis are free to rotate at any degree, then only $\{\Delta z_i - \Delta z_{i+1}\}_{i=0}^{n-2}$ affects the shape of piecewise linear function. As far as $\{\Delta z_i\}_{i=0}^{n-1}$ remains the same, any change in data magnitude z_i or knot spacing h_i only scales the piecewise linear curve, but does not change its shape.

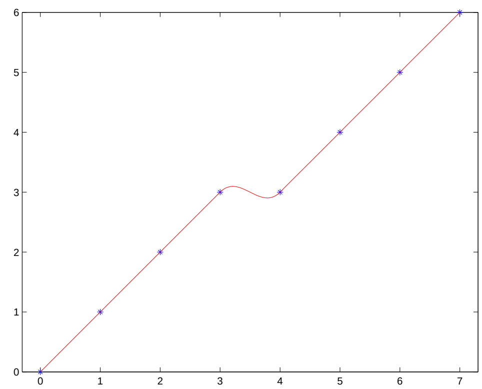

Figure 4.4: The cubic L_1 spline does not preserve monotonicity.

Cubic L_1 splines behave in a very similar way as piecewise linear interpolation functions. When $\{\Delta z_i\}_{i=0}^{n-1}$ keeps the same, the first derivatives q_i's do not change. However, since cubic L_1 splines are smoother than piecewise linear functions, their shape-preserving capabilities are weaker. If there is a change in the data magnitude z_i or knot spacing h_i, the shape of $S_i(x)$ will be changed. But this change only has local effect. The piecewise cubic functions over other sub-intervals are not affected. Therefore, cubic L_1 splines achieve a compromise between piecewise linear interpolation and C^1-smoothness. They preserve shape very well for multi-scale data, including data with abrupt changes in magnitude and spacing.

4.2.2 Monotonicity

Cubic L_1 splines do not preserve monotonicity in general. An example in shown in Figure 4.4, where the data set $\{(x_i, z_i)\}_{i=0}^{n}$ is given as $\{(0,0), (1,1), (2,2), (3,3), (4,3), (5,4), (6,5), (7,6)\}$.

4.2.3 Preserving linearity

We show in this section that if part of the given data set forms a straight line, then a cubic L_1 spline actually creates a linear function over this part. It is guaranteed that there is no nonphysical oscillation. This property can also be used to decompose the original interpolating problem into several small problems for parallel processing.

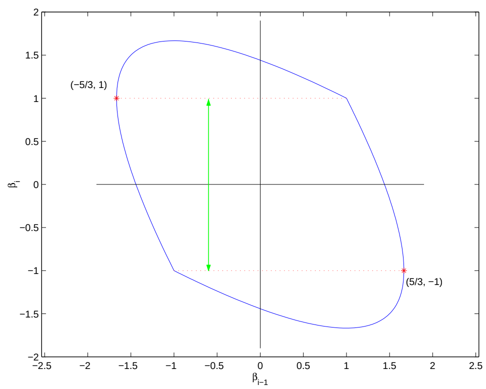

Figure 4.5: Illustration of the proof that L_1 spline preserves linearity.

Proposition 4.1 *If $\left(\beta_{i-1}^*, \beta_i^*\right)$ is an interior point of Ω, then $q_i = q_{i+1} = \Delta z_i$ and $S(x)$ is a linear function over $[x_i, x_{i+1}]$.*

Proof. If $\left(\beta_{i-1}^*, \beta_i^*\right)$ is an interior point of Ω, then $n_1\left(\beta_{i-1}^*, \beta_i^*\right) = n_2\left(\beta_{i-1}^*, \beta_i^*\right) = 0$. From model (T), $q_i = q_{i+1} = \Delta z_i$. This implies that $u_i = 0$ by (2.9), and $v_i = 0$ by (2.10). Hence, the L_1 spline over interval $[x_i, x_{i+1}]$ is a linear function

$$S_i(x) = p_i + \Delta z_i (x - x_i).$$

■

Theorem 4.2 (Cubic L_1 splines preserve linearity over more than three consecutive subintervals) *If there are four points (x_i, z_i), (x_{i+1}, z_{i+1}), (x_{i+2}, z_{i+2}), and (x_{i+3}, z_{i+3}) in the given data set lie on a straight line, then a cubic L_1 spline $S(x)$ preserves linearity over the intermediate subinterval $[x_{i+1}, x_{i+2}]$. Furthermore, $S(x)$ preserves linearity over the first subinterval $[x_i, x_{i+1}]$, if $\beta_{i-1}^* \neq \pm\frac{5}{3}$, and over the last subinterval $[x_{i+2}, x_{i+3}]$, if $\beta_{i+2}^* \neq \pm\frac{5}{3}$.*

Proof. For such a given data set, we know $\Delta z_i = \Delta z_{i+1} = \Delta z_{i+2}$. Hence, in the geometric dual problem (D), $b_i = b_{i+1} = 0$. Consequently, β_i^* and β_{i+1}^* can be any feasible values for a dual optimal solution,.

Assume $\beta_{i-1}^* \neq \pm\frac{5}{3}$ and $\beta_{i+2}^* \neq \pm\frac{5}{3}$. From the definition of Ω, as illustrated in Figure 4.5, it is not difficult to find β_i^* such that $\left(\beta_{i-1}^*, \beta_i^*\right)$ is an interior point of Ω and $|\beta_i^*| < 1$. Similarly, there exists a β_{i+1}^* such that $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ is an interior point of Ω and $\left|\beta_{i+1}^*\right| < 1$. Furthermore, for any $|\beta_i^*| < 1$ and $\left|\beta_{i+1}^*\right| < 1$, it is clear that $\left(\beta_i^*, \beta_{i+1}^*\right)$ is an interior point of Ω. Hence, we are able to find feasible β_i^* and β_{i+1}^* such that $\left(\beta_{i-1}^*, \beta_i^*\right)$, $\left(\beta_i^*, \beta_{i+1}^*\right)$, and $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ are all interior points of Ω. By Proposition 4.1, we have

$$q_i = q_{i+1} = q_{i+2} = q_{i+3} = \Delta z_i.$$

Therefore, the cubic L_1 spline $S(x)$ is a linear function over the three consecutive subintervals $[x_i, x_{i+3}]$.

If $\beta_{i-1}^* = \frac{5}{3}$ and $\beta_{i+2}^* \neq \pm\frac{5}{3}$, then the only feasible value of β_i^* is

$$\beta_i^* = -1.$$

In this case, $\left(\beta_{i-1}^*, \beta_i^*\right)$ is a boundary point of Ω, and $n_1\left(\beta_{i-1}^*, \beta_i^*\right) \geq 0$, $n_2\left(\beta_{i-1}^*, \beta_i^*\right) = 0$. From the dual-to-primal transformation model (T),

$$q_{i+1} = \Delta z_i + n_2\left(\beta_{i-1}^*, \beta_i^*\right) = \Delta z_i.$$

Furthermore, for any $\left|\beta_{i+1}^*\right| < 1$, $\left(-1, \beta_{i+1}^*\right)$ remains an interior point of Ω. Therefore,

$$q_{i+1} = q_{i+2} = q_{i+3} = \Delta z_i.$$

If $\beta_{i-1}^* = -\frac{5}{3}$ and $\beta_{i+2}^* \neq \pm\frac{5}{3}$, then the only feasible value of β_i^* is $\beta_i^* = 1$. The normal vector of Ω at $\left(\beta_{i-1}^*, \beta_i^*\right)$ satisfies $n_1\left(\beta_{i-1}^*, \beta_i^*\right) \leq 0$ and $n_2\left(\beta_{i-1}^*, \beta_i^*\right) = 0$. The same conclusion of $q_{i+1} = q_{i+2} = q_{i+3} = \Delta z_i$ holds.

The case of $\beta_{i-1}^* \neq \pm\frac{5}{3}$ and $\beta_{i+2}^* = \pm\frac{5}{3}$ can be proved following the same logic.

If $\beta_{i-1}^* = \frac{5}{3}$ and $\beta_{i+2}^* = -\frac{5}{3}$, then $\beta_i^* = -1$ and $\beta_{i+1}^* = 1$. It is easy to see that $\left(\beta_i^*, \beta_{i+1}^*\right) = (-1, 1)$ is an interior point of Ω. By Proposition 4.1, $S(x)$ is a straight line over the intermediate sub-interval $[x_{i+1}, x_{i+2}]$. Same for the case of $\beta_{i-1}^* = -\frac{5}{3}$ and $\beta_{i+2}^* = \frac{5}{3}$.

If $\beta_{i-1}^* = \frac{5}{3}$ and $\beta_{i+2}^* = \frac{5}{3}$, then $\left(\beta_i^*, \beta_{i+1}^*\right) = (-1, -1)$ is a corner point. From model (T), we have

$$\begin{aligned}
&q_{i+1} = \Delta z_i + n_2\left(\beta_{i-1}^*, \beta_i^*\right), \\
&q_{i+1} = \Delta z_{i+1} - n_1\left(\beta_i^*, \beta_{i+1}^*\right), \\
&n_2\left(\beta_{i-1}^*, \beta_i^*\right) = 0.
\end{aligned}$$

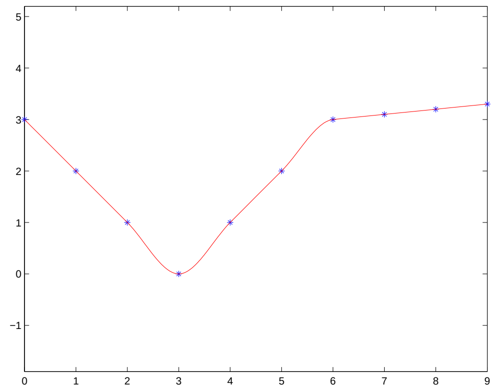

Figure 4.6: An example of the situation when some $\beta_i^* = \pm\frac{5}{3}$.

Solving this linear system achieves $n_1\left(\beta_i^*, \beta_{i+1}^*\right) = 0$, which further implies that $n_2\left(\beta_i^*, \beta_{i+1}^*\right) = 0$. Thus, $q_{i+2} = \Delta z_{i+1} + n_2\left(\beta_i^*, \beta_{i+1}^*\right) = \Delta z_{i+1} = q_{i+1}$, and $S(x)$ is a straight line over the intermediate sub-interval $[x_{i+1}, x_{i+2}]$. Same logic applies when $\beta_{i-1}^* = -\frac{5}{3}$ and $\beta_{i+2}^* = -\frac{5}{3}$. ■

For $\beta_{i-1}^* = \frac{5}{3}$, the only feasible value of $\left(\beta_{i-1}^*, \beta_i^*\right)$ is $\left(\beta_{i-1}^*, \beta_i^*\right) = \left(\frac{5}{3}, -1\right)$. So is the value of $\left(\beta_{i-2}^*, \beta_{i-1}^*\right)$ uniquely determined as $\left(-1, \frac{5}{3}\right)$. The normal vector of Ω at point $\left(\beta_{i-2}^*, \beta_{i-1}^*\right)$ is $n_1\left(\beta_{i-2}^*, \beta_{i-1}^*\right) = 0$ and $n_2\left(\beta_{i-2}^*, \beta_{i-1}^*\right) \geq 0$. By model (T), we have

$$q_{i-1} = \Delta z_{i-1} - n_1\left(\beta_{i-2}^*, \beta_{i-1}^*\right) = \Delta z_{i-1}, \tag{4.5}$$

$$q_i = \Delta z_{i-1} + n_2\left(\beta_{i-2}^*, \beta_{i-1}^*\right) \geq \Delta z_{i-1}. \tag{4.6}$$

Similarly, the normal vector of Ω at point $\left(\beta_{i-1}^*, \beta_i^*\right)$ is given by $n_1\left(\beta_{i-1}^*, \beta_i^*\right) \geq 0$ and $n_2\left(\beta_{i-1}^*, \beta_i^*\right) = 0$. Therefore,

$$q_i = \Delta z_i - n_1\left(\beta_{i-1}^*, \beta_i^*\right) \leq \Delta z_i, \tag{4.7}$$

$$q_{i+1} = \Delta z_i + n_2\left(\beta_{i-1}^*, \beta_i^*\right) = \Delta z_i. \tag{4.8}$$

Putting (4.5), (4.6), (4.7) and (4.8) together, it is clear to see that q_i is free to be any value between Δz_{i-1} and Δz_i, and it is independent of other q's. In the case of $\beta_{i-1}^* = -\frac{5}{3}$, we can derive the similar result that q_i can be any value between Δz_i and Δz_{i-1}. Therefore, for $\beta_{i-1}^* = \pm\frac{5}{3}$, we can always set $q_i = \Delta z_i$ to preserve linearity.

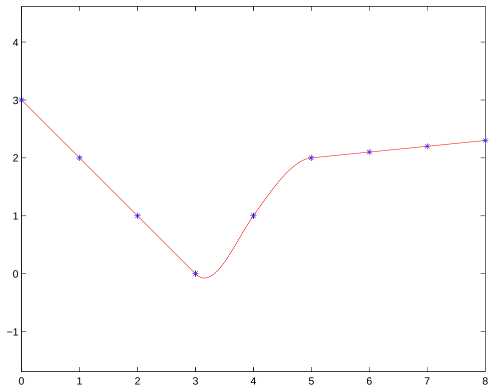

Figure 4.7: The cubic L_1 spline does not preserve linearity over less than three consecutive sub-intervals.

However, $\beta^*_{i-1} = \pm\frac{5}{3}$ usually occurs when x_i is at the intersection of two linear functions with different slopes. If q_i is set to be Δz_{i-1}, then the linearity on the left of x_i is preserved, but the linearity on the right is violated. The same problem occurs if we set $q_i = \Delta z_i$. In fact, only piecewise linear function is able to preserve the linearity of both sides of x_i in this case. But cubic L_1 splines preserve shape well in the sense that it fails to preserve linearity only over at most two sub-intervals around a intersection point. Furthermore, the given data is local convex (or concave) under this circumstance, and the properties discussed in the next section guarantee that any oscillation over the linear data will be eliminated.

When $\beta^*_{i-1} = \pm\frac{5}{3}$, from model (T), q_i is determined by the following one-dimensional optimization problem

$$\begin{array}{ll} \min & |q_i| \\ \text{s.t.} & \min\{\Delta z_{i-1}, \Delta z_i\} \leq q_i \leq \max\{\Delta z_{i-1}, \Delta z_i\}. \end{array}$$

Clearly, the solution is $q_i = \min\{|\Delta z_{i-1}|, |\Delta z_i|\}$, if Δz_{i-1} and Δz_i have the same sign. Otherwise, $q_i = 0$.

In summary, Theorem 4.2 guarantees that cubic L_1 splines preserve linearity over more than three consecutive sub-intervals. The only exception happens when x_i is the intersection point of two linear data sets. In this case, there does not exist any C^1-smooth interpolater to preserve linearity. Therefore, cubic L_1 splines achieve a good balance between preserving linearity and having C^1-smoothness.

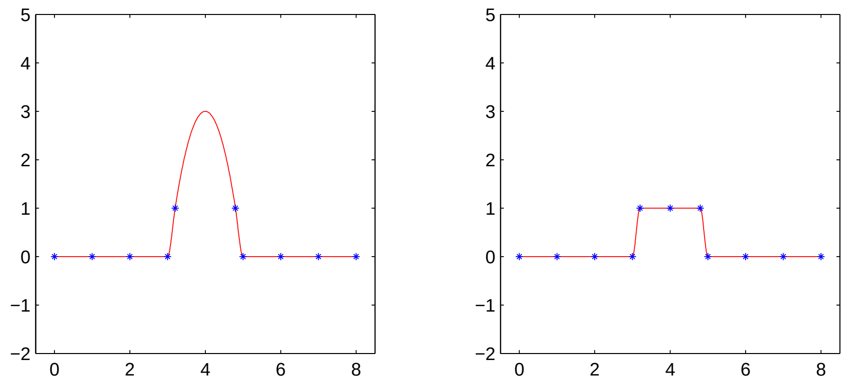

Figure 4.8: The cubic L_1 spline preserves shape better with a finer spacing.

An example is used to illustrate the conclusion of Theorem 4.2. As shown in Figure 4.6, the data set $\{(x_i, z_i)\}_{i=0}^{n}$ is given by $\{(0,3), (1,2), (2,1), (3,0), (4,1), (5,2), (6,3), (7,3.1), (8,3.2), (9,3.3)\}$. The dual optimal solution is $\beta_2^* = \frac{5}{3}$, $\beta_5^* = -\frac{5}{3}$ and $\beta_i^* = 0$, for $i \neq 2,5$. The data set is linear over intervals $[0,3]$, $[3,6]$, and $[6,9]$. At point $x = 3$, we can set $q_3 = -1$ to preserve linearity over $[0,3]$, or set $q_3 = 1$ to preserve linearity over $[3,6]$. After considering the regularization term, we have $q_3^* = 0$. At point $(6,3)$, the value of q_6 can be any number in $[0.1, 1]$. The regularization term gives the optimal value $q_6^* = 0.1$. The resulting cubic L_1 spline is the flattest one among infinite many possible L_1 splines.

The minimum number of consecutive intervals for cubic L_1 splines to preserve linearity is three. A counterexample is shown in Figure 4.7. The data set $\{(x_i, z_i)\}_{i=0}^{n}$ is given by $\{(0,3), (1,2), (2,1), (3,0), (4,1), (5,2), (6,2.1), (7,2.2), (8,2.3)\}$. As we have proved, the L_1 spline $S(x)$ is a linear function over $[0,3]$ and $[5,8]$. However, over two consecutive sub-intervals $[3,4]$ and $[4,5]$, even the piecewise linear interpolation function forms a straight line, $S(x)$ does not.

Theorem 4.2 has important implications in practice. If it is known *a priori* that one interval of the interpolated function is linear, then it is sufficient to pick only four knots in this interval. They are the two end points of the interval and two arbitrary intermediate points. Collecting more data would not improve the result and only raise the cost. On the other hand, for a general curve, introducing additional knots will force the cubic L_1 spline move close to the piecewise linear interpolation. An example is illustrated in Figure 4.8. For the left figure, the data set $\{(x_i, z_i)\}_{i=0}^{n}$ is given by $\{(0,0), (1,0), (2,0),$

$(3,0)$, $(3.2,1)$, $(4.8,1)$, $(5,0)$, $(6,0)$, $(7,0)$, $(8,0)\}$. While an additional knot $(4,1)$ is introduced, the L_1 spline changes drastically as shown in the right figure. Therefore, the finer the spacing is, the better the L_1 spline may preserve the shape.

4.2.4 Convexity/concavity and nonphysical oscillation

If the underlying function of the given data set is convex, i.e., the piecewise linear interpolation function is convex, then the cubic L_1 spline $S(x)$ preserves the shape in certain sense. If the divided difference $\{\Delta z_i\}_{i=0}^{n-2}$ do not change too much, then $S(x)$ is a convex function. Even when $S(x)$ is not convex, it still has some properties to guarantee that there is no nonphysical oscillation.

Theorem 4.3 *If $\Delta z_i \le \Delta z_{i+1} \le \Delta z_{i+2}$, then $q_{i+1} \le \Delta z_{i+1}$ and $q_{i+2} \ge \Delta z_{i+1}$.*

Proof. Since $q_{i+1} = \Delta z_{i+1} - n_1\left(\beta_i^*, \beta_{i+1}^*\right)$ and $q_{i+2} = \Delta z_{i+1} + n_2\left(\beta_i^*, \beta_{i+1}^*\right)$, the conclusion is equivalent to

$$n_1\left(\beta_i^*, \beta_{i+1}^*\right) \ge 0, \quad n_2\left(\beta_i^*, \beta_{i+1}^*\right) \ge 0, \tag{4.9}$$

for $\Delta z_i \le \Delta z_{i+1} \le \Delta z_{i+2}$.

If $\Delta z_i = \Delta z_{i+1} = \Delta z_{i+2}$, then this is the case discussed in Theorem 4.2. We have $q_{i+1} = q_{i+2} = \Delta z_{i+1}$. In the following, we assume that $\Delta z_i = \Delta z_{i+1}$ and $\Delta z_{i+1} = \Delta z_{i+2}$ do not hold at the same time.

If the respective dual optimal solution pair $\left(\beta_i^*, \beta_{i+1}^*\right)$ is an interior point of Ω, then Proposition 4.1 implies that $q_{i+1} = q_{i+2} = \Delta z_{i+1}$. The conclusion follows.

If $\left(\beta_i^*, \beta_{i+1}^*\right)$ is a boundary point of Ω and $\left(\beta_{i-1}^*, \beta_i^*\right)$ is an interior point of Ω, then Proposition 4.1 implies that $q_{i+1} = \Delta z_i$. From model (T), $q_{i+1} = \Delta z_{i+1} - n_1\left(\beta_i^*, \beta_{i+1}^*\right) = \Delta z_i$. Consequently,

$$n_1\left(\beta_i^*, \beta_{i+1}^*\right) = \Delta z_{i+1} - \Delta z_i \ge 0.$$

Next we need to show that under this circumstance, if $n_1\left(\beta_i^*, \beta_{i+1}^*\right) \ge 0$ then $n_2\left(\beta_i^*, \beta_{i+1}^*\right) \ge 0$. Assume $n_1\left(\beta_i^*, \beta_{i+1}^*\right) \ge 0$ and $n_2\left(\beta_i^*, \beta_{i+1}^*\right) < 0$. Then $\left(\beta_i^*, \beta_{i+1}^*\right)$ falls in section II of Ω, i.e., $-\frac{5}{3} \le \beta_{i+1}^* \le -1$. Consider the $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ pair. Since $-\frac{5}{3} \le \beta_{i+1}^* \le -1$, $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ can be either an interior point of

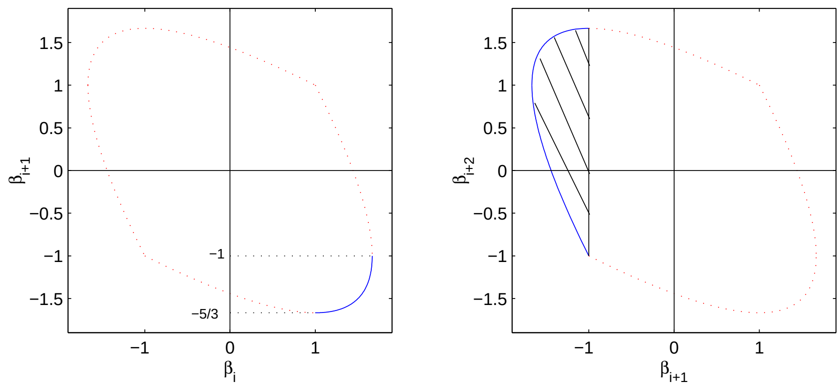

Figure 4.9: Illustration of the proof.

Ω, or a boundary point of Ω in section III or IV. In either case, $n_1\left(\beta_{i+1}^*, \beta_{i+2}^*\right) \leq 0$. The feasible set of $\left(\beta_i^*, \beta_{i+1}^*\right)$ and $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ are shown in Figure 4.9. Again, from the model (T),

$$q_{i+2} = \Delta z_{i+1} + n_2\left(\beta_i^*, \beta_{i+1}^*\right),$$

$$q_{i+2} = \Delta z_{i+2} - n_1\left(\beta_{i+1}^*, \beta_{i+2}^*\right),$$

which implies that

$$\Delta z_{i+2} - \Delta z_{i+1} = n_2\left(\beta_i^*, \beta_{i+1}^*\right) + n_1\left(\beta_{i+1}^*, \beta_{i+2}^*\right) < 0.$$

This contradicts the assumption of $\Delta z_{i+1} \leq \Delta z_{i+2}$. Therefore, we have $n_2\left(\beta_i^*, \beta_{i+1}^*\right) \geq 0$.

If $\left(\beta_i^*, \beta_{i+1}^*\right)$ is a boundary point of Ω and $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ is an interior point of Ω, then the same logic can show that (4.9) holds.

Next we assume that $\left(\beta_{i-1}^*, \beta_i^*\right)$, $\left(\beta_i^*, \beta_{i+1}^*\right)$, and $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ are all boundary points of Ω. As we have shown above, if one value of $n_1\left(\beta_i^*, \beta_{i+1}^*\right)$ and $n_2\left(\beta_i^*, \beta_{i+1}^*\right)$ is nonnegative, then the other one must be nonnegative also. Hence, we only need to consider the case that

$$n_1\left(\beta_i^*, \beta_{i+1}^*\right) \leq 0 \ \text{ and } \ n_2\left(\beta_i^*, \beta_{i+1}^*\right) \leq 0.$$

In this case, $\left(\beta_i^*, \beta_{i+1}^*\right)$ falls in the section III of Ω, i.e., $-\frac{5}{3} \leq \beta_i^* \leq 1, -\frac{5}{3} \leq \beta_{i+1}^* \leq 1$. Consider the $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ pair. Since $\Delta z_{i+2} - \Delta z_{i+1} = n_2\left(\beta_i^*, \beta_{i+1}^*\right) + n_1\left(\beta_{i+1}^*, \beta_{i+2}^*\right) \geq 0$ and $n_2\left(\beta_i^*, \beta_{i+1}^*\right) \leq 0$, we have $n_1\left(\beta_{i+1}^*, \beta_{i+2}^*\right) \geq 0$. Combining with $-\frac{5}{3} \leq \beta_{i+1}^* \leq 1$ results in $-1 \leq \beta_{i+1}^* \leq 1$. Similarly, by considering the $\left(\beta_{i-1}^*, \beta_i^*\right)$ pair we have $-1 \leq \beta_i^* \leq 1$. Since $\left(\beta_i^*, \beta_{i+1}^*\right)$ falls in the section III of Ω, the only possible value of $\left(\beta_i^*, \beta_{i+1}^*\right)$ is

$$\beta_i^* = \beta_{i+1}^* = -1.$$

However, this implies $\beta_{i+2}^* = \frac{5}{3}$ or -1, and therefore, $n_1\left(\beta_{i+1}^*, \beta_{i+2}^*\right) \leq 0$. Hence,

$$\Delta z_{i+2} - \Delta z_{i+1} = n_2\left(\beta_i^*, \beta_{i+1}^*\right) + n_1\left(\beta_{i+1}^*, \beta_{i+2}^*\right) \leq 0.$$

Similarly, we can get $\Delta z_{i+1} - \Delta z_i \leq 0$. This is true only when

$$\Delta z_i = \Delta z_{i+1} = \Delta z_{i+2},$$

and consequently,

$$n_1\left(\beta_i^*, \beta_{i+1}^*\right) = n_2\left(\beta_i^*, \beta_{i+1}^*\right) = 0.$$

Therefore, if $\left(\beta_{i-1}^*, \beta_i^*\right)$, $\left(\beta_i^*, \beta_{i+1}^*\right)$, and $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ are all boundary points of Ω, then $n_1\left(\beta_i^*, \beta_{i+1}^*\right) \geq 0$ and $n_2\left(\beta_i^*, \beta_{i+1}^*\right) \geq 0$. ■

Theorem 4.4 *If $\Delta z_i \leq \Delta z_{i+1} \leq \Delta z_{i+2}$, then the cubic L_1 spline $S(x)$ is convex over $[x_{i+1}, x_{i+2}]$ if and only if $q_{i+1} = q_{i+2} = \Delta z_{i+1}$ or $\beta_i^* = \beta_{i+1}^* = 1$.*

Proof. If $\Delta z_i \leq \Delta z_{i+1} \leq \Delta z_{i+2}$, then $n_1\left(\beta_i^*, \beta_{i+1}^*\right) \geq 0$ and $n_2\left(\beta_i^*, \beta_{i+1}^*\right) \geq 0$. For any $x \in [x_{i+1}, x_{i+2}]$,

$$S''(x) = u_{i+1} + v_{i+1}(x - x_{i+1})$$

is a linear function. Hence, $S''(x) \geq 0$ over $[x_{i+1}, x_{i+2}]$ if and only if

$$\begin{aligned} S''(x_{i+1}) &= u_{i+1} \\ &= \frac{2}{h_{i+1}}\left[2n_1\left(\beta_i^*, \beta_{i+1}^*\right) - n_2\left(\beta_i^*, \beta_{i+1}^*\right)\right] \geq 0, \end{aligned}$$

and

$$\begin{aligned} S''(x_{i+2}) &= u_{i+1} + v_{i+1}h_{i+1} \\ &= \frac{2}{h_{i+1}}\left[-n_1\left(\beta_i^*, \beta_{i+1}^*\right) + 2n_2\left(\beta_i^*, \beta_{i+1}^*\right)\right] \geq 0. \end{aligned}$$

If $n_1\left(\beta_i^*, \beta_{i+1}^*\right) = n_2\left(\beta_i^*, \beta_{i+1}^*\right) = 0$, then all these requirements are satisfied and $q_{i+1} = q_{i+2} = \Delta z_{i+1}$. Otherwise, the only value of $\left(\beta_i^*, \beta_{i+1}^*\right)$ satisfies all these requirements is $\beta_i^* = \beta_{i+1}^* = 1$. ■

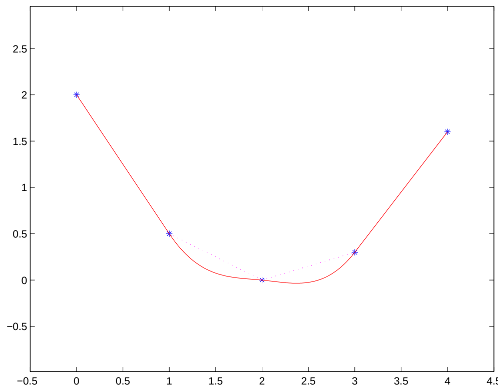

Figure 4.10: An example that the L_1 spline does not preserve convexity.

Figure 4.10 shows one example that the cubic L_1 spline $S(x)$ does not preserve convexity. The data set $\{(x_i, z_i)\}_{i=0}^{n}$ is given by $\{(0,2), (1,0.5), (2,0), (3,0.3), (4,1.6)\}$.

The dual optimal solution is $\beta_i^* = 1$ for the convex data only when the divided differences Δz_i do not increase too rapidly. For example, assuming $\Delta z_i \leq \Delta z_{i+1} \leq \Delta z_{i+2}$, consider the extreme case that both $\left(\beta_{i-1}^*, \beta_i^*\right)$ and $\left(\beta_{i+1}^*, \beta_{i+2}^*\right)$ are interior points of Ω. Then we only need to consider the convexity of cubic L_1 spline $S(x)$ over one sub-interval $[x_{i+1}, x_{i+2}]$. In this case, we have $q_{i+1} = \Delta z_i$ and $q_{i+2} = \Delta z_{i+2}$. From model (T), it is not difficult to derive that

$$\begin{aligned} n_1\left(\beta_i^*, \beta_{i+1}^*\right) &= \Delta z_{i+1} - \Delta z_i, \\ n_2\left(\beta_i^*, \beta_{i+1}^*\right) &= \Delta z_{i+2} - \Delta z_{i+1}. \end{aligned}$$

If the dual optimal solution is $\beta_i^* = \beta_{i+1}^* = 1$, then the normal vector at point $(1,1)$ should satisfy

$$\frac{1}{2} n_1\left(\beta_i^*, \beta_{i+1}^*\right) \leq n_2\left(\beta_i^*, \beta_{i+1}^*\right) \leq 2 n_1\left(\beta_i^*, \beta_{i+1}^*\right).$$

Putting them together, $S(x)$ preserves convexity if and only if

$$\frac{2}{3}\Delta z_i + \frac{1}{3}\Delta z_{i+2} \leq \Delta z_{i+1} \leq \frac{1}{3}\Delta z_i + \frac{2}{3}\Delta z_{i+2}.$$

When there are more sub-intervals need to be considered together, the condition for preserving convexity becomes highly complicated.

Even when $S(x)$ is not convex, it is guaranteed that $S(x)$ would never cross the piecewise linear interpolation function. Hence, $S(x)$ eliminates the possibility of nonphysical oscillation.

Theorem 4.5 *If $\Delta z_i \leq \Delta z_{i+1} \leq \Delta z_{i+2}$, then cubic L_1 spline $S(x)$ is bounded from above by the linear function $p_{i+1} + \Delta z_{i+1}(x - x_{i+1})$ over $[x_{i+1}, x_{i+2}]$.*

Proof. If $n_1(\beta_i^*, \beta_{i+1}^*) = n_2(\beta_i^*, \beta_{i+1}^*) = 0$, then $S(x) = p_{i+1} + \Delta z_{i+1}(x - x_{i+1})$ for any $x \in [x_{i+1}, x_{i+2}]$.

Assume that $n_1(\beta_i^*, \beta_{i+1}^*)$ and $n_2(\beta_i^*, \beta_{i+1}^*)$ are not zero at the same time. Let

$$\begin{aligned} W(x) &= S(x) - [p_{i+1} + \Delta z_{i+1}(x - x_{i+1})] \\ &= (q_{i+1} - \Delta z_{i+1})(x - x_{i+1}) + \frac{u_{i+1}}{2}(x - x_{i+1})^2 + \frac{v_{i+1}}{6}(x - x_{i+1})^3. \end{aligned}$$

Then $W(x_{i+1}) = W(x_{i+2}) = 0$, and

$$W'(x) = (q_{i+1} - \Delta z_{i+1}) + u_{i+1}(x - x_{i+1}) + \frac{v_{i+1}}{2}(x - x_{i+1})^2.$$

Therefore,

$$W'(x_{i+1}) = q_{i+1} - \Delta z_{i+1} = -n_1(\beta_i^*, \beta_{i+1}^*) \leq 0$$

$$\begin{aligned} W'(x_{i+2}) &= (q_{i+1} - \Delta z_{i+1}) + \frac{u_{i+1}}{2}h_{i+1} + \frac{v_{i+1}}{2}h_{i+1}^2 \\ &= q_{i+2} - \Delta z_{i+1} \\ &= n_2(\beta_i^*, \beta_{i+1}^*) \geq 0 \end{aligned}$$

Since $W'(x)$ is a non-constraint quadratic function, there must exist a unique root $\tilde{x} \in [x_{i+1}, x_{i+2}]$ such that, for any $x_{i+1} \leq x \leq \tilde{x}$, $W'(x) \leq 0$ and for any $\tilde{x} \leq x \leq x_{i+2}$, $W'(x) \geq 0$. Therefore, for any $x_{i+1} \leq x \leq \tilde{x}$, $W(x)$ is a monotonically non-increasing function, which implies $W(x) \leq W(x_{i+1}) = 0$. Similarly, for any $\tilde{x} \leq x \leq x_{i+2}$, $W(x)$ is a monotonically non-decreasing function with $W(x) \leq W(x_{i+2}) = 0$. Overall, for any $x \in [x_{i+1}, x_{i+2}]$,

$$S(x) \leq p_{i+1} + \Delta z_{i+1}(x - x_{i+1}).$$

∎

Any enforcement of local convexity tends to violate the conclusion of Theorem 4.5 and introduces non-physical oscillations.

Similar results can be obtained if the underlying function of a given data set is concave. We state these results as following without proofs.

Theorem 4.6 *If $\Delta z_i \geq \Delta z_{i+1} \geq \Delta z_{i+2}$, then $q_{i+1} \geq \Delta z_{i+1}$ and $q_{i+2} \leq \Delta z_{i+1}$.*

Theorem 4.7 *If $\Delta z_i \geq \Delta z_{i+1} \geq \Delta z_{i+2}$, then cubic L_1 spline $S(x)$ is concave over $[x_{i+1}, x_{i+2}]$ if and only if $q_{i+1} = q_{i+2} = \Delta z_{i+1}$ or $\beta_i^* = \beta_{i+1}^* = -1$.*

Theorem 4.8 *If $\Delta z_i \geq \Delta z_{i+1} \geq \Delta z_{i+2}$, then cubic L_1 spline $S(x)$ is bounded from below by the linear function $p_{i+1} + \Delta z_{i+1}(x - x_{i+1})$ over $[x_{i+1}, x_{i+2}]$.*

4.3 Conclusion

Geometric programming model provides a nice framework to perform theoretical analysis and develop efficient algorithms. This model has an essential advantage that it establishes the relationship between the coefficients of a cubic L_1 spline and an optimal solution of the geometric dual program. Following this work, we analyze the properties of cubic L_1 splines and explain why cubic L_1 splines are shape-preserving multi-scale interpolators for arbitrary data.

The first derivatives of cubic L_1 splines depend on the divided differences $\{\Delta z_i\}_{i=0}^{n-1}$ only. As far as $\{\Delta z_i\}_{i=0}^{n-1}$ remains unchanged, data magnitude and knot spacing only affect local scale but not the shape of a spline. Hence, cubic L_1 splines perform excellent for multi-scale data. This property is analogous to piecewise linear interpolation functions.

Cubic L_1 splines preserve linearity over more than three consecutive sub-intervals. The only exception is at the intersection of two linear segments with different slopes. In that circumstance, there does not exist any C^1-smooth interpolating function to be able to preserve linearity. Cubic L_1 splines are the best C^1-smooth interpolators to preserve linearity in the sense that they fails to do so over at most two sub-intervals around a intersection point. For any case, cubic L_1 splines eliminate oscillations over linear data.

When given data is convex/concave, cubic L_1 splines would never cross piecewise linear function, and therefore, eliminate any non-physical oscillation. However, they do not preserve convexity/concavity in general. The sufficient and necessary condition for preserving convexity/concavity is given in terms of the dual optimal solution.

In summary, cubic L_1 splines achieve a good balance between piecewise linear interpolation and C^1-smoothness. They preserve shape very well, in particular, in eliminating extraneous non-physical oscillation.

Chapter 5

Continuum-based algorithm for solving univariate cubic L_1 splines

In this chapter, a continuum-based algorithm for calculating the coefficients of univariate cubic L_1 splines is developed within the geometric programming framework. Section 5.2 gives the optimality theory and shows how to decompose the original problem. Section 5.3 develops the algorithm for solving the geometric dual program. Section 5.4 shows how to obtain the primal optimal solution. Section 5.5 discusses the convergence and complexity of the algorithm. Section 5.6 summarizes this chapter.

5.1 Notation

First we introduce some definitions and notation that will be used in the rest of the chapter.

Define

$$f(x) = |x| - \frac{\sqrt{40} - 2}{3}, \tag{5.1}$$

and

$$g(x, y) = \frac{1}{3}|x - y| - 1 + \frac{1}{4}(x + y)^2. \tag{5.2}$$

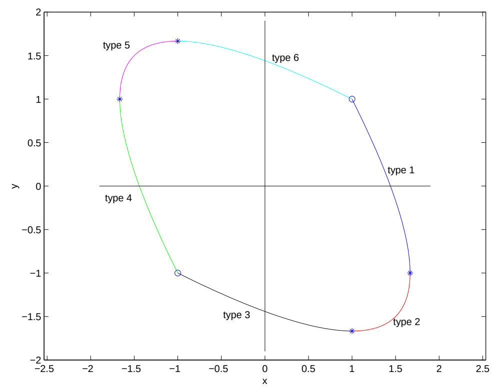

Figure 5.1: The feasible set of Ω and different regions of $\partial\Omega$.

Let Ω be the feasible set of a consecutive pair of dual variables (β_{i-1}, β_i), $i = 1, \dots n-2$,

$$\begin{aligned} \Omega &= \left\{ (x,y)^T \in R^2 \,|g(x,y) \leq 0 \right\} \\ &= \left\{ (x,y)^T \in R^2 \left| \frac{1}{3} |x-y| \leq 1 - \frac{1}{4}(x+y)^2 \right. \right\}. \end{aligned} \tag{5.3}$$

Ω is a convex set enclosed by the two quadratic curves

$$C_L = \left\{ (x,y)^T \in R^2 \left| -\frac{1}{3}(x-y) = 1 - \frac{1}{4}(x+y)^2 \right. \right\},$$

and

$$C_R = \left\{ (x,y)^T \in R^2 \left| \frac{1}{3}(x-y) = 1 - \frac{1}{4}(x+y)^2 \right. \right\}.$$

The feasible set Ω is illustrated in Figure 5.1.

Let $(n_1(x,y), n_2(x,y))^T$ be a normal vector of Ω at the point (x,y). A normal vector will be defined not only by the point (x,y), but also by non-negative Lagrange multipliers λ and μ. When dependence on λ and μ is of concern, a normal vector of Ω at the point (x,y) will be denoted $(n_1(x,y;\lambda,\mu), n_2(x,y;\lambda,\mu))^T$. The notation $n_1(x,y)$ and $n_1(x,y;\lambda,\mu)$ will be used interchangeably.

The points $\left(\frac{5}{3}, -1\right)$, $\left(1, -\frac{5}{3}\right)$, $\left(-\frac{5}{3}, 1\right)$ and $\left(-1, \frac{5}{3}\right)$ are called axis points. They divide the boundary of Ω into four sections. In each section, $n_1(x,y)$ maintains the same sign, as does $n_2(x,y)$. These four axis points are indicated by '*' in Figure 5.1.

The two points of intersection of C_L and C_R, namely, the points $(1,1)$ and $(-1,-1)$ are called corner points. At these two points, the two quadratic constraints of Ω are both active. The corner points are indicated by '$\circ$' in Figure 5.1.

The boundary of Ω is denoted by $\partial\Omega$, i.e.,

$$\partial\Omega = \left\{ (x, y)^T \in R^2 \,|g\,(x, y) = 0 \right\}.$$

By the axis points and the corner points, $\partial\Omega$ is divided into six "regions," namely,

$$\text{region 1:} \qquad \left\{ (x, y)^T \in \partial\Omega \,\middle|\, 1 \le x \le \frac{5}{3}, -1 \le y \le 1 \right\}, \tag{5.4}$$

$$\text{region 2:} \qquad \left\{ (x, y)^T \in \partial\Omega \,\middle|\, 1 \le x \le \frac{5}{3}, -\frac{5}{3} \le y \le -1 \right\}, \tag{5.5}$$

$$\text{region 3:} \qquad \left\{ (x, y)^T \in \partial\Omega \,\middle|\, -1 \le x \le 1, -\frac{5}{3} \le y \le -1 \right\}, \tag{5.6}$$

$$\text{region 4:} \qquad \left\{ (x, y)^T \in \partial\Omega \,\middle|\, -\frac{5}{3} \le x \le -1, -1 \le y \le 1 \right\}, \tag{5.7}$$

$$\text{region 5:} \qquad \left\{ (x, y)^T \in \partial\Omega \,\middle|\, -\frac{5}{3} \le x \le -1, 1 \le y \le \frac{5}{3} \right\}, \tag{5.8}$$

and

$$\text{region 6:} \qquad \left\{ (x, y)^T \in \partial\Omega \,\middle|\, -1 \le x \le 1, 1 \le y \le \frac{5}{3} \right\}. \tag{5.9}$$

Two dual variables $\boldsymbol{\beta}^1$ and $\boldsymbol{\beta}^2$ are said to be in the same region if, for any $i = 1, \ldots, n-2$, $\left(\beta_{i-1}^1, \beta_i^1\right)$ and $\left(\beta_{i-1}^2, \beta_i^2\right)$ are either all interior points of Ω or are in the same region.

A normal vector for any point (x, y) in regions 1, 2, and 3 except the corner points is calculated by

$$\begin{pmatrix} n_1\,(x, y; \lambda, \mu) \\ \\ n_2\,(x, y; \lambda, \mu) \end{pmatrix} = \lambda \begin{pmatrix} x + y + \frac{2}{3} \\ \\ x + y - \frac{2}{3} \end{pmatrix}, \quad \lambda \ge 0. \tag{5.10}$$

A normal vector for any point (x, y) in regions 4, 5, and 6 except the corner points is calculated by

$$\begin{pmatrix} n_1\,(x, y; \lambda, \mu) \\ \\ n_2\,(x, y; \lambda, \mu) \end{pmatrix} = \lambda \begin{pmatrix} x + y - \frac{2}{3} \\ \\ x + y + \frac{2}{3} \end{pmatrix}, \quad \lambda \ge 0. \tag{5.11}$$

The normal vectors at the corner point $(1, 1)$ form the cone

$$\begin{pmatrix} n_1\,(1, 1; \lambda, \mu) \\ \\ n_2\,(1, 1; \lambda, \mu) \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ \\ 1 \end{pmatrix}, \quad \lambda, \mu \ge 0. \tag{5.12}$$

Similarly, the normal vectors at the corner point $(-1,-1)$ form the cone

$$\begin{pmatrix} n_1(-1,-1;\lambda,\mu) \\ \\ n_2(-1,-1;\lambda,\mu) \end{pmatrix} = \lambda \begin{pmatrix} -1 \\ \\ -2 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ \\ -1 \end{pmatrix}, \quad \lambda,\mu \geq 0. \tag{5.13}$$

A general normal vector is a linear combination of two basic normal vectors

$$n_l(x,y;\lambda,\mu) = \lambda \cdot n_l(x,y;1,0) + \mu \cdot n_l(x,y;0,1), \quad l = 1,2. \tag{5.14}$$

with coefficients λ and μ.

The assignment operation is denoted by "$\leftarrow$," that is, "$x \leftarrow y$" means "assign the value of y to x." The statement "$x \rightleftharpoons y$" means "interchange the values of x and y." Logical equality is denoted by "$==$".

5.2 Optimality theory and partition of the problem

5.2.1 Optimality theory

The dual problem is to minimize a linear function over a bounded closed convex set. It is well known that

Theorem 5.1 (Existence) *The optimal solution of the dual problem (D) exists.*

An optimality condition for the dual optimal solution can be derived from the definition of the normal cone.

Theorem 5.2 *A dual feasible solution $\boldsymbol{\beta}^*$ is optimal if and only if $\langle B, \boldsymbol{\beta}^* - \boldsymbol{\beta}\rangle \leq 0$ for any feasible solution $\boldsymbol{\beta}$, i.e., $\boldsymbol{\beta}^*$ is a dual optimal solution if and only if $-B$ is in the normal cone of the feasible set at $\boldsymbol{\beta}^*$.*

The normal cone of the feasible set can be constructed by applying the following theorem

Theorem 5.3 ([24] Corollary 23.8.1) *Let $C_1,\ldots,C_m$ be convex sets in R^n whose relative interiors have a point in common. Then the normal cone to $C_1 \cap \cdots \cap C_m$ at any given point x is $K_1 + \cdots + K_m$, where K_i is the normal cone to C_i at x.*

Consider the dual problem (D). Obviously the zero vector is an interior point of the feasible set. For $i = 1, \ldots, n-2$, the set

$$\left\{\boldsymbol{\beta} \in R^{n-1} \mid g\left(\beta_{i-1}, \beta_i\right) \le 0\right\}$$

is a convex set. As can easily be verified, its normal vector at any $\boldsymbol{\beta}$ is

$$\left(\underbrace{0, \ldots, 0}_{i-2}, n_1\left(\beta_{i-1}, \beta_i\right), n_2\left(\beta_{i-1}, \beta_i\right), \underbrace{0, \ldots, 0}_{n-i-1}\right),$$

where $n_1\left(\beta_{i-1}, \beta_i\right)$ and $n_2\left(\beta_{i-1}, \beta_i\right)$ are defined by (5.10)-(5.13). For $i = 0$, the set

$$\left\{\boldsymbol{\beta} \in R^{n-1} \mid f\left(\beta_0\right) \le 0\right\}$$

is convex. A normal vector of this set is

$$\left(\lambda_0, \underbrace{0, \ldots, 0}_{n-2}\right),$$

with $\lambda_0 \ge 0$ and $\lambda_0 \cdot f\left(\beta_0\right) = 0$. For the last constraint, the set

$$\left\{\boldsymbol{\beta} \in R^{n-1} \mid f\left(\beta_{n-2}\right) \le 0\right\}$$

is convex. A normal vector of this set is

$$\left(\underbrace{0, \ldots, 0}_{n-2}, \lambda_{n-1}\right),$$

with $\lambda_{n-1} \ge 0$ and $\lambda_{n-1} \cdot f\left(\beta_{n-2}\right) = 0$.

The following optimality condition is a direct conclusion of the statements above.

Theorem 5.4 *A dual feasible solution $\boldsymbol{\beta}^*$ is optimal if and only if there exist vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ such that*

$$1) - b_0 = \lambda_0 + n_1\left(\beta_0^*, \beta_1^*; \lambda_1, \mu_1\right), \quad \lambda_0 \cdot f\left(\beta_0^*\right) = 0, \tag{5.15}$$

$$2) - b_i = n_2\left(\beta_{i-1}^*, \beta_i^*; \lambda_i, \mu_i\right) + n_1\left(\beta_i^*, \beta_{i+1}^*; \lambda_{i+1}, \mu_{i+1}\right), \quad i = 1, \ldots, n-3, \tag{5.16}$$

$$3) - b_{n-2} = \lambda_{n-1} + n_2\left(\beta_{n-3}^*, \beta_{n-2}^*; \lambda_{n-2}, \mu_{n-2}\right), \quad \lambda_{n-1} \cdot f\left(\beta_{n-2}^*\right) = 0, \tag{5.17}$$

$$4) \;\; \lambda_i \ge 0, \quad \mu_i \ge 0, \quad i = 0, \ldots, n-1. \tag{5.18}$$

The vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are called Lagrange multipliers.

5.2.2 Segmentation and data perturbation

Consider two consecutive dual variables β_{i-1} and β_i. In the dual problem (D), these two variables are related by the constraint $g(\beta_{i-1}, \beta_i) \leq 0$, $i = 1, \ldots, n-2$. If $g(\beta_{i-1}, \beta_i) = 0$, then (β_{i-1}, β_i) is a boundary point of Ω. Otherwise, (β_{i-1}, β_i) is an interior point of Ω.

Let us think of the dual variable vector $\boldsymbol{\beta}$ as a sequence of real numbers $\beta_0, \ldots, \beta_{n-2}$. This sequence of numbers can be partitioned into subsequences in the following manner. Two consecutive variables β_{i-1} and β_i will be in the same subsequence if and only if (β_{i-1}, β_i) is a boundary point of Ω, i.e., $g(\beta_{i-1}, \beta_i) = 0$. Let the total number of subsequences be denoted by M. The mth subsequence, $1 \leq m \leq M$, is a sequence of numbers $\beta_{k_m}, \ldots, \beta_{k_m + l_m}$, where $l_m \geq 0$; l_m is the "length" of this subsequence. If β_{i-1} and β_i are in the same subsequence, i.e., $k_m + 1 \leq i \leq k_m + l_m$ for some $1 \leq m \leq M$, then $g(\beta_{i-1}, \beta_i) = 0$. If β_{i-1} and β_i are not in the same subsequence, i.e., $i = k_m$ or $i - 1 = k_m + l_m$ for some $1 \leq m \leq M$, then $g(\beta_{i-1}, \beta_i) < 0$. We call subsequences defined in this manner "segments of the dual variable $\boldsymbol{\beta}$." Let σ_m denote the mth segment, i.e., $\sigma_m = \{\beta_{k_m}, \ldots, \beta_{k_m + l_m}\}$, $1 \leq m \leq M$. The decomposition of dual variables into segments is called a segmentation and is denoted by $\Sigma = \{\sigma_1, \ldots, \sigma_M\}$.

Let $\boldsymbol{\beta}$ be an arbitrary dual feasible solution. It is possible that there is a β_l such that both (β_{l-1}, β_l) and (β_l, β_{l+1}) are interior points of Ω. Hence, β_l does not belong to any segment of $\boldsymbol{\beta}$. In this case, we say that the segmentation of $\boldsymbol{\beta}$ is invalid. A segmentation $\Sigma = \{\sigma_1, \ldots, \sigma_m\}$ of $\boldsymbol{\beta}$ is said to be valid if only if each β_i, $i = 0, \ldots, n-2$, belongs to one and only one σ_m, $1 \leq m \leq M$. The following corollary of Theorem 5.4 guarantees that there exists at least one dual optimal solution, the segmentation of which is valid.

Corollary 5.5 *There exists an optimal solution $\boldsymbol{\beta}^*$ such that at least one of any two consecutive pairs $\left(\beta_{i-1}^*, \beta_i^*\right)$ and $\left(\beta_i^*, \beta_{i+1}^*\right)$ must be a boundary point of Ω. Consequently, $\boldsymbol{\beta}^*$ always has a valid segmentation.*

Proof. Assume that there exist $\left(\beta_{i-1}^*, \beta_i^*\right)$ and $\left(\beta_i^*, \beta_{i+1}^*\right)$ that are both interior points of Ω.

Then

$$n_1\left(\beta_{i-1}^*, \beta_i^*\right) = n_2\left(\beta_{i-1}^*, \beta_i^*\right) = 0$$

and

$$n_1\left(\beta_i^*, \beta_{i+1}^*\right) = n_2\left(\beta_i^*, \beta_{i+1}^*\right) = 0.$$

Hence, by Theorem 5.4,

$$b_i = -n_2\left(\beta_{i-1}^*, \beta_i^*\right) - n_1\left(\beta_i^*, \beta_{i+1}^*\right) = 0.$$

Since the coefficient of β_i in the objective function is zero, β_i^* can take any feasible value without changing the objective function value. Therefore, β_i^* can be increased or decreased until either $\left(\beta_{i-1}^*, \beta_i^*\right)$ or $\left(\beta_i^*, \beta_{i+1}^*\right)$ reaches $\partial\Omega$. ■

In the rest of this chapter, we consider only dual feasible solutions with valid segmentations.

Segments in which some (β_{j-1}, β_j) is a corner point or an axis point have a special property. Let (β_{j-1}, β_j) be a corner or axis point. If $\beta_{j-1} = 1$, then, by definition (5.2), the feasible value of β_j can only be 1 or $-\frac{5}{3}$. If $\beta_{j-1} = -\frac{5}{3}$, then β_j can only be 1. We observe similar results for other axis points and corner points. In summary, we have

Theorem 5.6 *Let $\boldsymbol{\beta}$ be a dual feasible solution with a valid segmentation. If, in a segment, there exists a variable pair that is an axis point, then every variable pair in this segment is either an axis point or a corner point. If, in a segment, there exists a corner point but no axis points, then every variable in this segment must have the same value and this value can be only* 1 *or* -1.

Let $\boldsymbol{\beta}^*$ be a dual optimal solution. We will decompose the dual problem (D) into separate subproblems (DS) for the variables $\sigma_m = \left\{\beta_{k_m}^*, \ldots, \beta_{k_m+l_m}^*\right\}$, within each segment:

$$\text{(DS)} \quad \begin{cases} \min\limits_{\beta_i} & \sum\limits_{i=k_m}^{k_m+l_m} b_i\beta_i \\ \text{s.t.} & g\left(\beta_{i-1}, \beta_i\right) \leq 0, \quad i = k_m+1, \ldots, k_m+l_m. \end{cases}$$

For the first segment σ_1, the boundary condition $f\left(\beta_0\right) \leq 0$ needs to be added to subproblem (DS). For the last segment σ_M, the boundary condition $f\left(\beta_{n-2}\right) \leq 0$ needs to be added to (DS). The following result is a direct consequence of Theorem 5.4.

Corollary 5.7 *If $\boldsymbol{\beta}^*$ is a dual optimal solution with segmentation $\Sigma = \{\sigma_1, \ldots, \sigma_M\}$, then $\left(\beta^*_{k_m}, \ldots, \beta^*_{k_m+l_m}\right)^T$ is an optimal solution of subproblem (DS) for any $m = 1, \ldots, M$. Let $\boldsymbol{\beta}$ be a dual feasible solution with valid segmentation $\Sigma = \{\sigma_1, \ldots, \sigma_M\}$. If, for each segment σ_m, $(\beta_{k_m}, \ldots, \beta_{k_m+l_m})^T$ is an optimal solution of subproblem (DS), then $\boldsymbol{\beta}$ is an optimal solution of the original dual problem (D).*

The primal solution can be segmented in the same manner. Let $\boldsymbol{\beta}^*$ be a dual optimal solution with segmentation $\Sigma = \{\sigma_1, \ldots, \sigma_M\}$. If, in one segment $\beta^*_{k_m}, \ldots, \beta^*_{k_m+l_m}$, the pair $\left(\beta^*_{k_m-1}, \beta^*_{k_m}\right)$ is an interior point of Ω, then $n_1\left(\beta^*_{k_m-1}, \beta^*_{k_m}\right) = 0$ and $n_2\left(\beta^*_{k_m-1}, \beta^*_{k_m}\right) = 0$. Therefore, from model (T), we know that

$$q_{k_m+1} = \Delta z_{k_m} + n_2\left(\beta^*_{k_m-1}, \beta^*_{k_m}\right) = \Delta z_{k_m}.$$

Similarly,

$$q_{k_m+l_m+1} = \Delta z_{k_m+l_m+1}.$$

Hence, the corresponding segment of the primal optimal solution $q^*_{k_m+1}, \ldots, q^*_{k_m+l_m+1}$ is an optimal solution to the subproblem

$$
(\mathrm{TS}) \quad \left\{
\begin{array}{lrcl}
\displaystyle\min_{q_i, \lambda_i, \mu_i} & \displaystyle\sum_{i=k_m+1}^{k_m+l_m+1} |q_i| & & \\
\text{s.t.} & q_{k_m+1} & = & \Delta z_{k_m}, \\
& q_i & = & \Delta z_i - n_1\left(\beta^*_{i-1}, \beta^*_i; \lambda_i, \mu_i\right), \ i = k_m+1, \ldots, k_m+l_m, \\
& q_{i+1} & = & \Delta z_i + n_2\left(\beta^*_{i-1}, \beta^*_i; \lambda_i, \mu_i\right), \ i = k_m+1, \ldots, k_m+l_m, \\
& q_{k_m+l_m+1} & = & \Delta z_{k_m+l_m+1}, \\
& \lambda_i, \mu_i & \geq & 0.
\end{array}
\right.
$$

Cubic L_1 splines thus have the advantage of enabling parallel computation by decomposing the dual problem into several independent small-scale subproblems (DS) to be solved in parallel. Then the primal optimal solution can be achieved by solving small-scale problems (TS), also in parallel. Perturbations of a function value z_i that do not change the segmentation affect the dual and primal solutions only in the segment containing the perturbed point. The optimal solutions in other segments remain unchanged. Thus, to update a cubic L_1 spline, it is sufficient to solve small-scale problems (DS) and (TS).

5.3 Algorithm for solving the dual problem

5.3.1 Outline of the algorithm

A general convex program can be formulated as

$$\begin{cases} \min\limits_{\mathbf{x}\in R^n} & h_0(\mathbf{x}) \\ \text{s.t.} & h_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots k. \end{cases}$$

where $h_i(\mathbf{x})$, $i = 0, 1, \ldots k$, is a convex function. Define the active constraint set to be

$$\Gamma = \{i \in \{1, \ldots k\} \,|h_i(\mathbf{x}) = 0\}.$$

Let ρ_i be the Lagrange multiplier corresponding to the constraint $h_i(\mathbf{x}) \leq 0$.

The so called active set algorithm for solving the convex program is as follows.

Algorithm 5.1 (Active set algorithm [11])

Step 1: Choose the initial point $\mathbf{x}^{(0)}$ and the initial active set $\Gamma^{(0)}$. $k \leftarrow 0$.

Step 2: Find the optimal solution $\mathbf{x}^{(k+1)}$ and Lagrange multipliers $\boldsymbol{\rho}^{(k+1)}$ under the condition that all constraints in the set $\Gamma^{(k)}$ are active.

Step 3: $\Gamma^{(k+1)} \leftarrow \Gamma^{(k)}$. If $j \notin \Gamma^{(k+1)}$ and $h_j\left(\mathbf{x}^{(k+1)}\right) == 0$, then $\Gamma^{(k+1)} \leftarrow \Gamma^{(k+1)} + \{j\}$. If $j \in \Gamma^{(k+1)}$ and $\rho_j^{(k+1)} < 0$, then $\Gamma^{(k+1)} \leftarrow \Gamma^{(k+1)} - \{j\}$.

Step 4: If $\Gamma^{(k+1)} == \Gamma^{(k)}$, then stop. Otherwise, $k \leftarrow k+1$ and go to Step 2.

The special structure of (D) suggests that there are several advantages in applying this algorithm to solve the dual problem (D). First, for the dual problem, the active constraint set of a feasible solution $\boldsymbol{\beta}$ is defined by the segmentation of $\boldsymbol{\beta}$. By the definition of segmentation, two consecutive variables β_{i-1} and β_i are in the same segment if and only if $g(\beta_{i-1}, \beta_i) = 0$, i.e., the constraint $g(\beta_{i-1}, \beta_i) \leq 0$ is active. Second, by Corollary 5.7, computation of Step 1 in Algorithm 5.1 can be decomposed into several independent small-scale subproblems. The optimal solution in one segment is not affected by the optimal solutions in other segments. Third, when the active set changes, i.e., when the segmentation changes, only the solutions of the subproblems in the changed segments need to be recalculated. In Algorithm 5.1, after obtaining a new solution $\mathbf{x}^{(k+1)}$, one constraint may become active, in which case its index needs to be added to the active set $\Gamma^{(k+1)}$. In the dual problem (D), this happens only when $\beta_{i-1}^{(k)}$ and $\beta_i^{(k)}$ belong to two different segments and $g\left(\beta_{i-1}^{(k+1)}, \beta_i^{(k+1)}\right) = 0$. Hence, these two segments need to be merged into one. Another kind of change is the splitting of one segment. That happens when two variables $\beta_{i-1}^{(k)}$ and $\beta_i^{(k)}$ belong to the same segment and, after one iteration, the Lagrange multiplier corresponding to the active constraint $g\left(\beta_{i-1}^{(k)}, \beta_i^{(k)}\right) = 0$ is negative. Assume $\beta_i^{(k)} \in \sigma_m^{(k)}$. Then $\sigma_m^{(k+1)}$ contains only from $\beta_{k_m}^{(k+1)}$ up to $\beta_{i-1}^{(k+1)}$ and the new $\sigma_{m+1}^{(k+1)}$ consists of $\beta_i^{(k+1)}, \ldots, \beta_{k_m+l_m}^{(k+1)}$. There may be some segments that neither merge with another segment nor split into two. The optimal solution of such a segment remains unchanged and need not be calculated again.

We now apply the general active set algorithm to our problem and modify it to make best use of these advantages.

Algorithm 5.2 (Cubic L_1 spline active set algorithm)

Step 1: Choose the initial dual feasible solution $\boldsymbol{\beta}$ and get its segmentation $\Sigma = \{\sigma_1, \ldots, \sigma_M\}$. $m \leftarrow 1$.

Step 2: Find the optimal solution for segment σ_m.

Step 3.1: If σ_m merges with σ_{m-1}, then $m \leftarrow m - 1$, $M \leftarrow M - 1$ and go to Step 2.

Step 3.2: If σ_m merges with σ_{m+1}, then $M \leftarrow M - 1$ and go to Step 2.

Step 4: Calculate Lagrange multipliers λ_i and μ_i within σ_m. If there is a $\lambda_{i_0} < 0$ or $\mu_{i_0} < 0$, then split σ_m at $(\beta_{i_0-1}, \beta_{i_0})$, make the assignment $M \leftarrow M + 1$ and go to Step 2.

Step 5: $m \leftarrow m + 1$. If $m \leq M$, then go to Step 2. Otherwise, stop.

Return $\boldsymbol{\beta}$.

The details of Step 2 are discussed in Section 5.3.2. Section 5.3.3 is dedicated to the details of Step 4. Finally, Section 5.3.4 presents a method for choosing the starting point $\boldsymbol{\beta}$.

5.3.2 Optimization in one segment

In this section we discuss how to find the optimal solution in one segment. For convenience of notation, assume that the current segment consists of $k + 1$ variables $\beta_0, \ldots, \beta_k$. Then, we need to solve the problem

$$\text{(DE)} \quad \begin{cases} \min\limits_{\boldsymbol{\beta}} \sum\limits_{i=0}^{k} b_i \beta_i \\ \text{s.t.} \quad g\left(\beta_{i-1}, \beta_i\right) = 0, \quad i = 1, \ldots, k. \end{cases}$$

Its feasible set is

$$\Theta = \left\{ \boldsymbol{\beta} \in R^{k+1} \,|\, g\left(\beta_{i-1}, \beta_i\right) = 0, \quad i = 1, \ldots, k \right\}.$$

For any $\boldsymbol{\beta} \in \Theta$, $g\left(\beta_{i-1}, \beta_i\right) = 0$, i.e., $\left(\beta_{i-1}, \beta_i\right)$ is an interior point of Ω. Hence, for any i, $i = 1, 2, \ldots, k$, $\left(\beta_{i-1}, \beta_i\right)$ must be located in one of the six regions of $\partial\Omega$. Once it is known which region $\left(\beta_{i-1}, \beta_i\right)$ is in, one can write β_i as a function of β_{i-1} using (5.2). Specifically, if $\left(\beta_{i-1}, \beta_i\right)$ is a point in region 1, then

$$\beta_i\left(\beta_{i-1}\right) = -\beta_{i-1} + \frac{2}{3} + \frac{\sqrt{40 - 24\beta_{i-1}}}{3}; \tag{5.19}$$

if $\left(\beta_{i-1}, \beta_i\right)$ is a point in region 2 or 3, then

$$\beta_i\left(\beta_{i-1}\right) = -\beta_{i-1} + \frac{2}{3} - \frac{\sqrt{40 - 24\beta_{i-1}}}{3}; \tag{5.20}$$

if $\left(\beta_{i-1}, \beta_i\right)$ is a point in region 4, then

$$\beta_i\left(\beta_{i-1}\right) = -\beta_{i-1} - \frac{2}{3} - \frac{\sqrt{40 + 24\beta_{i-1}}}{3}; \tag{5.21}$$

if $\left(\beta_{i-1}, \beta_i\right)$ is a point in region 5 or 6, then

$$\beta_i\left(\beta_{i-1}\right) = -\beta_{i-1} - \frac{2}{3} + \frac{\sqrt{40 + 24\beta_{i-1}}}{3}. \tag{5.22}$$

As a consequence, once we known which region $\left(\beta_{i-1}, \beta_i\right)$ is in, $i = 1, 2, \ldots, k$, the values of β_i, $i = 1, 2, \ldots, k$, can be recursively calculated based on the value of β_0. Hence, a feasible solution β has only one degree of freedom, which can be taken to be β_0. Since all functions defined by (5.19)-(5.22) are strictly monotone, one could express all of the the β_i, $i = 0, 1, \ldots, k$, as functions of any one of the β_i. In summary, we have the following theorem

Theorem 5.8 *Given a priori knowledge of the regions of* $\left(\beta_{i-1}, \beta_i\right)$, $i = 1, 2, \ldots, k$, *a feasible solution* $\boldsymbol{\beta}$ *is uniquely determined by the value of any one variable* β_i, $i \in \{0, \ldots, k\}$.

In what follows, we use β_0 to develop the algorithm. In practice, it would be more efficient for parallel computation to choose the middle point $\beta_{\left[\frac{k}{2}\right]}$ and calculate backward from $\beta_{\left[\frac{k}{2}\right]}$ to β_0 and forward from $\beta_{\left[\frac{k}{2}\right]}$ to β_k simultaneously.

Let *region* [1..k] be an integer-valued array where the value of the ith element, *region*[i], indicates which region the point (β_{i-1}, β_i) is in. Let $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)^T$ be a feasible solution generated from a given value of β_0 and the array *region*[1..k]. As a consequence of Theorem 5.8, the objective function of problem (DE) is actually a one-dimensional function of β_0, which we denote as

$$\Phi\left(\beta_0\right) = \sum_{i=0}^{k} b_i \beta_i\left(\beta_0\right). \tag{5.23}$$

Let $\tilde{\boldsymbol{\beta}} = \left(\tilde{\beta}_0, \ldots, \tilde{\beta}_k\right)^T$ be a feasible solution calculated using $\tilde{\beta}_0 = \beta_0 + \delta$, where δ is a small real number. In order to determine whether $\boldsymbol{\beta}$ is a local minimizer of (DE), we need to compare the values of $\Phi\left(\tilde{\beta}_0\right)$ and $\Phi\left(\beta_0\right)$. The vector $\boldsymbol{\beta}$ is a local minimizer only if there exists a positive number $\delta_0 > 0$ such that

$$\Phi\left(\beta_0\right) \leq \Phi\left(\beta_0 + \delta\right), \quad \forall \left|\delta\right| \leq \delta_0. \tag{5.24}$$

If the current solution $\boldsymbol{\beta}$ is not optimal, we need to know in which direction to move β_0 in order to decrease $\Phi\left(\beta_0\right)$. There are three different cases to be considered.

Case 1: $\boldsymbol{\beta}$ contains no corner points and no axis points

Since every (β_{i-1}, β_i), $i = 1, \ldots k$, is neither a corner point nor an axis point, there exists a small $\delta_0 > 0$, such that for any $|\delta| \leq \delta_0$, $\left(\tilde{\beta}_{i-1}, \tilde{\beta}_i\right)$ is in the same region as (β_{i-1}, β_i). From (5.19)-(5.22), each β_i is a differentiable function of β_{i-1} and, therefore, is a differentiable function of β_0. As a consequence, $\Phi\left(\beta_0\right)$ is a differentiable function and $\Phi\left(\beta_0 + \delta\right)$ can be approximated by the differential

$$\Phi\left(\beta_0 + \delta\right) \approx \Phi\left(\beta_0\right) + \Phi'\left(\beta_0\right) \cdot \delta.$$

If $\Phi'\left(\beta_0\right) > 0$, then decreasing β_0 will decrease the value of $\Phi\left(\beta_0\right)$. Similarly, if $\Phi'\left(\beta_0\right) < 0$, then increasing β_0 will decrease the value of $\Phi\left(\beta_0\right)$. Therefore, $\boldsymbol{\beta}$ is a local minimizer of (DE) only if

$$\Phi'\left(\beta_0\right) = 0. \tag{5.25}$$

The derivative $\Phi'(\beta_0)$ is given by

$$\begin{aligned}
\Phi'(\beta_0) &= \sum_{i=0}^{k} b_i \beta_i'(\beta_0) \\
&= b_0 + \sum_{i=1}^{k} b_i \beta_i'(\beta_{i-1}) \beta_{i-1}'(\beta_0) \\
&= b_0 + \sum_{i=1}^{k} b_i \left(-\frac{n_1(\beta_{i-1}, \beta_i; \lambda_i, \mu_i)}{n_2(\beta_{i-1}, \beta_i; \lambda_i, \mu_i)} \right) \beta_{i-1}'(\beta_0) . \qquad (5.26)
\end{aligned}$$

The normal vector $(n_1(\beta_{i-1}, \beta_i), n_2(\beta_{i-1}, \beta_i))^T$ is easily calculated using (5.10) and (5.11). Furthermore, since any nonzero normal vector results in the same derivative $\beta_i'(\beta_{i-1})$ in (5.26), we specially choose $\lambda_i = 1$ and $\mu_i = 0$ in (5.10) and (5.11). Formula (5.26) can be implemented using the following recursive algorithm.

Algorithm 5.3 (Calculate $\Phi'(\beta_0)$ when no corner points and no axis points)

Let $\boldsymbol{\beta}$ and $region[1..k]$ be given.

Step 0: $\Phi'(\beta_0) \leftarrow b_0$, $d \leftarrow 1$, $i \leftarrow 1$.

For $i = 1..k$

Step 1: $d \leftarrow -\frac{n_1(\beta_{i-1}, \beta_i; 1, 0)}{n_2(\beta_{i-1}, \beta_i; 1, 0)} \cdot d$.

Step 2: $\Phi'(\beta_0) \leftarrow \Phi'(\beta_0) + b_i \cdot d$.

end For

Case 2: $\boldsymbol{\beta}$ contains corner points but no axis points

Assume that at least one of (β_{i-1}, β_i) is a corner point and no (β_{i-1}, β_i) is an axis point. In this case, for a small δ, the feasible solution $\tilde{\boldsymbol{\beta}} = \left(\tilde{\beta}_0, \ldots, \tilde{\beta}_k\right)$ is uniquely determined by the value of $\tilde{\beta}_0 = \beta_0 + \delta$. However, unlike the previous case, there are two possible regions in which each corner point locates.

For example, assume $\beta_0 = \beta_1 = \beta_2 = 1$. The corner point (β_0, β_1) can be considered to be in either region 1 or region 6. If δ is a small positive number, then $\left(\tilde{\beta}_0, \tilde{\beta}_1\right)$ is in region 1 and $\tilde{\beta}_1$ is calculated by (5.4). This is illustrated in the left of Figure 5.2, where the corner point (β_0, β_1) is indicated by 'o' and $\left(\tilde{\beta}_0, \tilde{\beta}_1\right)$ is indicated by '+'. If (β_0, β_1) is considered to be in region 1, the same region as $\left(\tilde{\beta}_0, \tilde{\beta}_1\right)$,

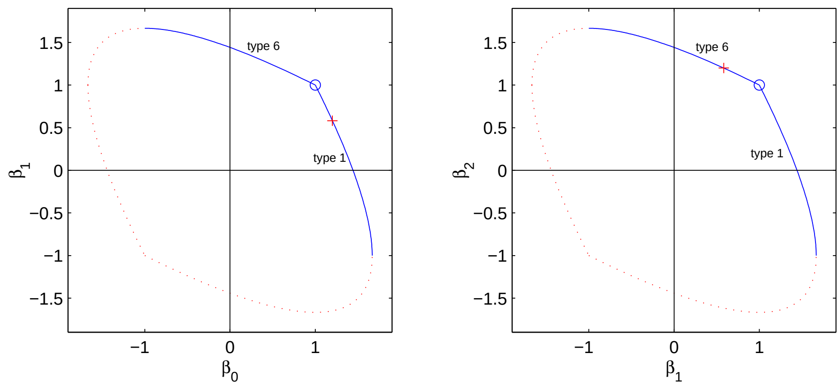

Figure 5.2: Example of the case with corner points.

then the right-hand derivative of $\beta_1(\beta_0)$ exists. Since $\tilde{\beta}_1 < 1$, $\left(\tilde{\beta}_1, \tilde{\beta}_2\right)$ must be in region 6 and $\tilde{\beta}_2$ is calculated by (5.9). The points (β_1, β_2) and $\left(\tilde{\beta}_1, \tilde{\beta}_2\right)$ are indicated in the right side of Figure 5.2 by 'o' and '+', respectively. If (β_1, β_2) is considered to be in region 6, then the left-hand derivative of $\beta_2(\beta_1)$ exists. On the other hand, if δ is a small negative number rather than a small positive number, then we conclude by analogous reasoning that (β_0, β_1) must be in region 6 and the left-hand derivative of $\beta_1(\beta_0)$ exists; and (β_1, β_2) must be in region 1 and the right-hand derivative of $\beta_2(\beta_1)$ exists.

When any one of the (β_{i-1}, β_i) is a corner point, $\Phi(\beta_0)$ is no longer differentiable. A corner point (β_{i-1}, β_i) can be in either one of two regions of $\partial\Omega$. If we consider the points of these two regions together, then β_i is a well-defined continuous function of β_{i-1}. Hence, given a small value of δ, regardless of whether δ is positive or negative, $\tilde{\boldsymbol{\beta}}$ is uniquely determined by $\tilde{\beta}_0 = \beta_0 + \delta$. If $\boldsymbol{\beta}$ and $\tilde{\boldsymbol{\beta}}$ are in the same region, then the array $region[1..k]$ of $\boldsymbol{\beta}$ is uniquely determined and both one-sided derivatives of $\beta_i(\beta_{i-1})$ exist. Consequently, even though $\Phi(\beta_0)$ is not differentiable, its two one-sided derivatives exist.

Let $\Phi'_-(\beta_0)$ and $\Phi'_+(\beta_0)$ be the left-hand and right-hand derivative, respectively. When $\delta = \tilde{\beta}_0 - \beta_0 > 0$, we have

$$\begin{aligned}\Phi(\beta_0 + \delta) &= \Phi(\beta_0 + sign(\delta) \cdot |\delta|) \\ &\approx \Phi(\beta_0) + \Phi'_+(\beta_0) \cdot \delta \\ &= \Phi(\beta_0) + \Phi'_+(\beta_0) \cdot sign(\delta) \cdot |\delta|, \quad \delta > 0.\end{aligned}$$

When $\delta = \tilde{\beta}_0 - \beta_0 < 0$, then

$$\begin{aligned}
\Phi(\beta_0 + \delta) &= \Phi(\beta_0 + sign(\delta) \cdot |\delta|) \\
&\approx \Phi(\beta_0) + \Phi'_-(\beta_0) \cdot \delta \\
&= \Phi(\beta_0) + \Phi'_-(\beta_0) \cdot sign(\delta) \cdot |\delta|, \quad \delta < 0.
\end{aligned}$$

Define a variable Δ to indicate the sign of $\delta = \tilde{\beta}_0 - \beta_0$ by

$$\Delta = \begin{cases} 1, & \text{if } \tilde{\beta}_0 - \beta_0 > 0 \\ -1, & \text{if } \tilde{\beta}_0 - \beta_0 < 0 \end{cases}. \tag{5.27}$$

Let

$$D(\Delta) = \lim_{|\delta| \to 0} \frac{\Phi(\beta_0 + \Delta \cdot |\delta|) - \Phi(\beta_0)}{|\delta|}. \tag{5.28}$$

Obviously, $D(1) = \Phi'_+(\beta_0)$, $D(-1) = -\Phi'_-(\beta_0)$, and

$$\Phi(\beta_0 + \delta) \approx \Phi(\beta_0) + D(sign(\delta)) \cdot |\delta|.$$

The current solution $\boldsymbol{\beta}$ is locally optimal only if (5.24) is satisfied. Condition (5.24) is equivalent to $D(1) \geq 0$ and $D(-1) \geq 0$.

If $D(1) > 0$ and $D(-1) < 0$, then $\Phi'_+(\beta_0) > 0$ and $\Phi'_-(\beta_0) > 0$. In this situation, β_0 needs to be decreased to reduce the value of $\Phi(\beta_0)$. On the other hand, if $D(1) < 0$ and $D(-1) > 0$, then $\Phi'_+(\beta_0) < 0$ and $\Phi'_-(\beta_0) < 0$. In this situation, we should increase β_0 to find the optimal solution.

In summary, if one defines

$$D^* = \min_{\Delta \in \{-1,1\}} D(\Delta) \tag{5.29}$$

and

$$\Delta^* = \arg \min_{\Delta \in \{-1,1\}} D(\Delta), \tag{5.30}$$

then the current solution $\boldsymbol{\beta}$ is locally optimal only if

$$D^* \geq 0. \tag{5.31}$$

If $D^* < 0$, we should move β_0 along the direction indicated by Δ^* to reduce the objective function value.

To calculate $D(\Delta)$ is essentially to calculate the one-sided derivatives $\Phi'_-(\beta_0)$ and $\Phi'_+(\beta_0)$. Algorithm 5.3 can be modified to do the work. In Step 1, the derivative $\beta'_i(\beta_{i-1}) = -\frac{n_1(\beta_{i-1},\beta_i;1,0)}{n_2(\beta_{i-1},\beta_i;1,0)}$

Table 5.1: $T\_Move[0..1, 1..6]$

region of (β_{i-1}, β_i)	1	2	3	4	5	6
$0\left(\tilde{\beta}_{i-1} < \beta_{i-1}\right)$	1	0	1	1	0	1
$1\left(\tilde{\beta}_{i-1} > \beta_{i-1}\right)$	0	1	0	0	1	0

needs to be replaced by the one-sided derivatives of $\beta_i\,(\beta_{i-1})$. If $(\beta_{i-1}, \beta_i) = (1, 1)$, then the left-hand derivative of $\beta_i\,(\beta_{i-1})$ is $-\frac{1}{2}$, and the right-hand derivative is -2. If $(\beta_{i-1}, \beta_i) = (-1, -1)$, then the left-hand and right-hand derivatives of $\beta_i\,(\beta_{i-1})$ are -2 and $-\frac{1}{2}$, respectively.

Furthermore, we need to compare the values of β_i and $\tilde{\beta}_i$ to decide the correct value of $region[i]$, and therefore, to decide which one-sided derivative of $\beta_i\,(\beta_{i-1})$ is the correct one to use. In order to be able to do that, it is necessary to keep track of the sign of $\tilde{\beta}_i - \beta_i$ for all $i = 0, \ldots, k$. Define a variable $iMove$ by

$$iMove = \begin{cases} 0, & \text{if } \tilde{\beta}_i < \beta_i \\ 1, & \text{if } \tilde{\beta}_i > \beta_i \end{cases}.$$

The $iMove$ value of β_i is uniquely determined by $region[i]$ and the $iMove$ value of β_{i-1}. Their relationship is summarized in the table $T\_Move[0..1, 1..6]$.

The first column of T_Move is read as follows: If (β_{i-1}, β_i) is a point in region 1 and $\tilde{\beta}_{i-1} < \beta_{i-1}$, then $\tilde{\beta}_i > \beta_i$. If (β_{i-1}, β_i) is a point in region 1 and $\tilde{\beta}_{i-1} > \beta_{i-1}$, then $\tilde{\beta}_i < \beta_i$. This conclusion is easy to verify since region 1 is defined by a monotonically decreasing function (5.19). The other columns correspond to the other 5 regions and can be explained in a similar manner. Hence, once the comparison of $\tilde{\beta}_0$ and β_0 is given, the comparison of $\tilde{\beta}_i$ and β_i is obtained recursively by

$$iMove \leftarrow T\_Move[iMove, region[i]].$$

The following algorithm summarizes the calculation of $D\,(\Delta)$ given a value of Δ,

Algorithm 5.4 (Calculate $D\,(\Delta)$)

Let $\boldsymbol{\beta}$, Δ and $region[1..k]$ be given.

Step 0: $D\,(\Delta) \leftarrow b_0 \cdot \Delta$ and $d \leftarrow \Delta$.

For $i = 1..k$

Step 1: $iMove \leftarrow T\_Move[iMove, region[i]]$.

Step 2: If $(\beta_{i-1}, \beta_i) == (1,1)$ then

If $iMove == 0$, then $region[i] \leftarrow 1$, $d \leftarrow -2 \cdot d$.

Otherwise, $region[i] \leftarrow 6$, $d \leftarrow -\frac{1}{2} \cdot d$.

end If

end If

If $(\beta_{i-1}, \beta_i) == (-1,-1)$ then

If $iMove == 0$, then $region[i] \leftarrow 3$, $d \leftarrow -\frac{1}{2} \cdot d$.

Otherwise $region[i] \leftarrow 4$, $d \leftarrow -2 \cdot d$.

end If

Otherwise, $d \leftarrow -\frac{n_1(\beta_{i-1},\beta_i;1,0)}{n_2(\beta_{i-1},\beta_i;1,0)} \cdot d$.

end If

Step 3: $D(\Delta) \leftarrow D(\Delta) + b_i \cdot d$.

end For

return $D(\Delta)$, $region[1..k]$, and d.

Case 3: $\boldsymbol{\beta}$ contains axis points

This is the case that there is at least one point (β_{i-1}, β_i) that is an axis point. In this case, for a small δ, the feasible solution $\tilde{\boldsymbol{\beta}} = \left(\tilde{\beta}_0, \ldots, \tilde{\beta}_k\right)$ corresponding to $\tilde{\beta}_0 = \beta_0 + \delta$ has multiple values.

An example is illustrated in Figure 5.3. In this example, both $(\beta_0, \beta_1) = \left(1, -\frac{5}{3}\right)$ and $(\beta_1, \beta_2) = \left(-\frac{5}{3}, 1\right)$ are axis points. They are indicated by '*' in Figure 5.3. For small $\delta > 0$, $\left(\tilde{\beta}_0, \tilde{\beta}_1\right)$ is a point in region 2 and is indicated by '+' in the left graph. For this $\tilde{\beta}_1$, the value of $\tilde{\beta}_2$ has two choices. The point $\left(\tilde{\beta}_1, \tilde{\beta}_2\right)$ indicated in the right graph by '+' can be a point in either region 4 or region 5. Hence, the value of $\tilde{\boldsymbol{\beta}}$ is not determined by $\tilde{\beta}_0$ alone, but also depends on the region of $\left(\tilde{\beta}_1, \tilde{\beta}_2\right)$ or depends on the sign of $\tilde{\beta}_1 - \beta_1$. If $\left(\tilde{\beta}_1, \tilde{\beta}_2\right)$ is a point in region 5 , then from (5.20) and (5.22), $\tilde{\beta}_2$ is an identity

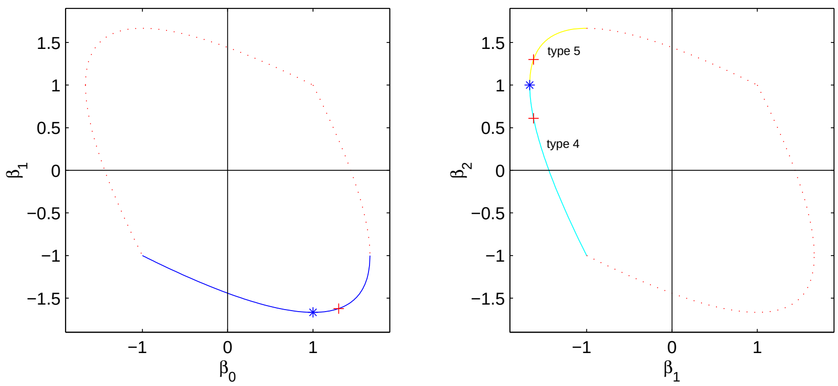

Figure 5.3: Example of the case with axis points.

function of $\tilde{\beta}_0$:

$$\tilde{\beta}_2\left(\tilde{\beta}_0\right) = \tilde{\beta}_2\left(\tilde{\beta}_1\left(\tilde{\beta}_0\right)\right) = \tilde{\beta}_0.$$

Hence,

$$\tilde{\beta}_2 - \beta_2 = \tilde{\beta}_0 - \beta_0.$$

If $\left(\tilde{\beta}_1, \tilde{\beta}_2\right)$ is a point in region 4 , then from (5.20) and (5.21), $\tilde{\beta}_2$ can be expressed as the following function of $\tilde{\beta}_0$:

$$\tilde{\beta}_2\left(\tilde{\beta}_0\right) = \tilde{\beta}_2\left(\tilde{\beta}_1\left(\tilde{\beta}_0\right)\right) = \tilde{\beta}_0 + \frac{2\sqrt{40 - 24\tilde{\beta}_0}}{3} - \frac{8}{3}.$$

Hence,

$$\left.\frac{d\tilde{\beta}_2}{d\tilde{\beta}_0}\right|_{\tilde{\beta}_0=1} = -1,$$

and, therefore,

$$\tilde{\beta}_2 - \beta_2 \approx -\left(\tilde{\beta}_0 - \beta_0\right).$$

It is always true that, at the axis point $(\beta_0, \beta_1) = \left(1, -\frac{5}{3}\right)$,

$$\frac{d\beta_1}{d\beta_0} = 0\,.$$

Hence, for small $\delta = \tilde{\beta}_0 - \beta_0$, the change $\tilde{\beta}_1 - \beta_1$ is infinitesimal and can be ignored. The above results together with results for the other similar situations can be summarized as follows.

Theorem 5.9 *If $\beta_l = \pm\frac{5}{3}$, then, for small δ,*

$$\tilde{\beta}_{l+1} - \beta_{l+1} \approx sign\left(\tilde{\beta}_{l+1} - \beta_{l+1}\right) \cdot sign\left(\tilde{\beta}_{l-1} - \beta_{l-1}\right) \cdot \left(\tilde{\beta}_{l-1} - \beta_{l-1}\right) \tag{5.32}$$

and

$$\sum_{i=0}^{k} b_i\left(\tilde{\beta}_i - \beta_i\right) \approx \sum_{\substack{i=0\\ i\neq l}}^{k} b_i\left(\tilde{\beta}_i - \beta_i\right). \tag{5.33}$$

Assume that there are N variables $\beta_{l_1}, \ldots, \beta_{l_N}$ that take on the value $\pm\frac{5}{3}$. These N variables cut the current segment into $N+1$ parts. For convenience of notation, let $l_0 = -1$ and $l_{N+1} = k+1$. Then part j contains variables $\beta_{l_j+1}, \ldots, \beta_{l_{j+1}}$, $j = 0, \ldots, N$. From Theorem 5.9, the variables $\beta_{l_1}, \ldots, \beta_{l_N}$ can be neglected in the objective function. Hence, for each part j, we only need consider variables $\beta_{l_j+1}, \ldots, \beta_{l_{j+1}-1}$. Since there is no axis point in part j, once the value of $\tilde{\beta}_{l_j+1}$ is given, all values $\tilde{\beta}_{l_j+1}, \ldots, \tilde{\beta}_{l_{j+1}-1}$ are, as previously discussed, uniquely determined. However, the value of $\tilde{\beta}_{l_j+1}$ depends on the sign of $\tilde{\beta}_{l_j+1} - \beta_{l_j+1}$ and is not be determined by $\tilde{\beta}_0$ alone. For $j = 0, \ldots, N$, let

$$\Delta_j = \begin{cases} 1, & \text{if } \tilde{\beta}_{l_j+1} - \beta_{l_j+1} > 0 \\ -1, & \text{if } \tilde{\beta}_{l_j+1} - \beta_{l_j+1} < 0 \end{cases},$$

$$\Delta = (\Delta_0, \ldots, \Delta_N)^T$$

and

$$\Phi_j\left(\beta_{l_j+1}\right) = \sum_{i=l_j+1}^{l_{j+1}-1} b_i \beta_i\left(\beta_{l_j+1}\right).$$

Analogous to what we did previously, define a function D of Δ by

$$D(\Delta) = \lim_{|\delta|\to 0} \frac{\sum_{j=0}^{N}\left\{\Phi_j\left(\beta_{l_j+1} + \Delta_j \cdot \left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|\right) - \Phi_j\left(\beta_{l_j+1}\right)\right\}}{|\delta|}. \tag{5.34}$$

The objective function at $\tilde{\boldsymbol{\beta}}$ can be approximated as follows:

$$\sum_{j=0}^{N} \Phi_j\left(\tilde{\beta}_{l_j+1}\right) \approx \sum_{j=0}^{N} \Phi_j\left(\beta_{l_j+1}\right) + D\left(sign\left(\tilde{\beta}_0 - \beta_0\right), \ldots, sign\left(\tilde{\beta}_{l_N} - \beta_{l_N}\right)\right) \cdot |\delta|.$$

Define

$$D^* = \min_{\Delta \in \{-1,1\}^{N+1}} D(\Delta) \tag{5.35}$$

and

$$\Delta^* = \arg\min_{\Delta \in \{-1,1\}^{N+1}} D(\Delta). \tag{5.36}$$

If we move the current solution $\boldsymbol{\beta}$ along the direction indicated by Δ^*, then the objective function $\sum_{j=0}^{N} \Phi_j \left(\beta_{l_j+1}\right) = \sum_{i=0}^{k} b_i\beta_i$ decreases most rapidly, that is, at the rate D^*. The current solution $\boldsymbol{\beta}$ is locally optimal only if

$$D^* \geq 0. \tag{5.37}$$

If $D^* < 0$, then, to reduce the objective function value, we need to move β_0, $\beta_{l_1}, \ldots, \beta_{l_N}$ along the direction indicated by Δ^*.

All functions defined by (5.19)-(5.22) are continuous. Hence, any finite composition of them is also a continuous function. As a consequence, $\left|\tilde{\beta}_0 - \beta_0\right|$ going to zero implies that $\left|\tilde{\beta}_i - \beta_i\right|$ goes to zero. Equation (5.34) can be simplified in the following manner:

$$\begin{aligned}
& D(\Delta) \\
= \quad & \lim_{|\delta|\to 0} \sum_{j=0}^{N} \frac{\left\{\Phi_j\left(\beta_{l_j+1} + \Delta_j \cdot \left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|\right) - \Phi_j\left(\beta_{l_j+1}\right)\right\}}{|\delta|} \\
= \quad & \lim_{|\delta|\to 0} \frac{\Phi_0\left(\beta_0 + \Delta_0 \cdot \left|\tilde{\beta}_0 - \beta_0\right|\right) - \Phi_0\left(\beta_0\right)}{|\delta|} \\
& + \sum_{j=1}^{N} \lim_{\left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|\to 0} \frac{\Phi_j\left(\beta_{l_j+1} + \Delta_j \cdot \left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|\right) - \Phi_j\left(\beta_{l_j+1}\right)}{\left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|} \\
& \cdot \lim_{|\delta|\to 0} \frac{\left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|}{\left|\tilde{\beta}_{l_j-1} - \beta_{l_j-1}\right|} \cdot \lim_{|\delta|\to 0} \frac{\left|\tilde{\beta}_{l_j-1} - \beta_{l_j-1}\right|}{|\delta|} \\
= \quad & D_0\left(\Delta_0\right) + \sum_{j=1}^{N} D_j\left(\Delta_j\right) \cdot \lim_{|\delta|\to 0} \frac{\left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|}{\left|\tilde{\beta}_{l_j-1} - \beta_{l_j-1}\right|} \cdot \lim_{|\delta|\to 0} \frac{\left|\tilde{\beta}_{l_j-1} - \beta_{l_j-1}\right|}{|\delta|},
\end{aligned}$$

where $D_j\left(\Delta_j\right)$ is (5.28) applied to $\beta_{l_j+1}, \ldots, \beta_{l_{j+1}-1}$ and can be calculated by Algorithm 5.4. By Theorem 5.9, the term $\lim_{|\delta|\to 0} \frac{\left|\tilde{\beta}_{l_j+1} - \beta_{l_j+1}\right|}{\left|\tilde{\beta}_{l_j-1} - \beta_{l_j-1}\right|}$ is always 1. The term $\lim_{|\delta|\to 0} \frac{\left|\tilde{\beta}_{l_j-1} - \beta_{l_j-1}\right|}{|\delta|}$ is the absolute value of one-sided derivative of $\beta_{l_j-1}\left(\beta_0\right)$. Its value depends on β_0 and $\Delta_0, \ldots, \Delta_{j-1}$. If we denote the one-sided derivative of $\beta_{l_j-1}(\beta_0)$ by $\beta'_{l_j-1}\left(\beta_0;\Delta_{l_{j-1}},\ldots,\Delta_0\right)$, then

$$\begin{aligned}
&\lim_{|\delta|\to 0}\frac{\left|\tilde{\beta}_{l_j-1}-\beta_{l_j-1}\right|}{|\delta|}\\
=\;&\left|\beta'_{l_j-1}\left(\beta_0;\Delta_0,\ldots,\Delta_{j-1}\right)\right|\\
=\;&\left|\beta'_{l_j-1}\left(\beta_{l_{j-1}+1};\Delta_{j-1}\right)\right|\cdot\left|\beta'_{l_{j-1}-1}\left(\beta_0;\Delta_0,\ldots,\Delta_{j-2}\right)\right|\\
=\;&\cdots\\
=\;&\left|\beta'_{l_j-1}\left(\beta_{l_{j-1}+1};\Delta_{j-1}\right)\right|\cdot\left|\beta'_{l_{j-1}-1}\left(\beta_{l_{j-2}+1};\Delta_{j-2}\right)\right|\cdots\left|\beta'_{l_1-1}\left(\beta_0;\Delta_0\right)\right|.
\end{aligned}$$

Problem (5.35) is a discrete optimization problem over all possible combinations of Δ. In the worst situation, there are corner points in every part. Consequently, $D(\Delta)$ may have up to 2^{N+1} different values. It is impractical to calculate all these values. Dynamic programming [5, 18] is applied to solve the problem.

Let stage j be the subproblem of (5.35) and (5.36) containing only variables from part j to part N. Let $R_j(\Delta_j)$ be the return of stage j under the condition that the value of Δ_j is given. Then

$$\begin{aligned}
&R_j(\Delta_j)\\
=\;&\min_{\Delta_{j+1},\ldots,\Delta_N} D\left(\Delta_j,\Delta_{j+1},\ldots,\Delta_N\right)\\
=\;&\min_{\Delta_{j+1},\ldots,\Delta_N}\left\{D_j(\Delta_j)+\sum_{r=j+1}^{N}D_r(\Delta_r)\cdot\left|\beta'_{l_r-1}\left(\beta_{l_j+1};\Delta_j,\ldots,\Delta_{r-1}\right)\right|\right\}\\
=\;&\min_{\Delta_{j+1},\ldots,\Delta_N}\left\{D_j(\Delta_j)+D_{j+1}(\Delta_{j+1})\cdot\left|\beta'_{l_{j+1}-1}\left(\beta_{l_j+1};\Delta_j\right)\right|\right.\\
&\left.+\sum_{r=j+2}^{N}D_r(\Delta_r)\cdot\left|\beta'_{l_r-1}\left(\beta_{l_{j-1}+1};\Delta_{j+1},\ldots,\Delta_{r-1}\right)\right|\cdot\left|\beta'_{l_{j+1}-1}\left(\beta_{l_j+1};\Delta_j\right)\right|\right\}\\
=\;&D_j(\Delta_j)+\min_{\Delta_{j+1},\ldots,\Delta_N}\left\{D_{j+1}(\Delta_{j+1})\right.\\
&\left.+\sum_{r=j+2}^{N}D_r(\Delta_r)\cdot\left|\beta'_{l_r-1}\left(\beta_{l_{j-1}+1};\Delta_{j+1},\ldots,\Delta_{r-1}\right)\right|\right\}\cdot\left|\beta'_{l_{j+1}-1}\left(\beta_{l_j+1};\Delta_j\right)\right|\\
=\;&D_j(\Delta_j)+\min\left\{R_{j+1}(1),R_{j+1}(-1)\right\}\cdot\left|\beta'_{l_{j+1}-1}\left(\beta_{l_j+1};\Delta_j\right)\right|.
\end{aligned}$$

Therefore, all values of $R_j(\Delta_j)$ can be calculated from stage N backward to stage 0. Obviously,

$$D^*=\min\left\{R_0(1),R_0(-1)\right\}.$$

The procedure of calculating (5.35) and (5.36) is summarized in the following algorithm. In the algorithm, an intermediate array $R[0..1, 0..N]$ is used to store the return of each stage. The value of $R[\Delta_j, j]$ is assigned to $D_j(\Delta_j)$ first and then is changed to be $R_j(\Delta_j)$. The variable D^* always keeps the value of $\min R_j(\Delta_j)$ and Δ_j^* is assigned the value of $\arg\min R_j(\Delta_j)$.

Another intermediate array $region\_t[0..1, 0..k]$ is used to keep the region of each variable pair temporarily. At the end of the algorithm, the correct region is assigned according to the value of Δ^*.

Algorithm 5.5 (Calculate D^*)

Let $\boldsymbol{\beta}$ and $region[1..k]$ be given.

Step 1: Find $\beta_{l_1}, \ldots, \beta_{l_N}$ that divide $\boldsymbol{\beta}$ into $N+1$ parts. $D^* \leftarrow 0$.

Step 2: $j \leftarrow N$.

Step 3: $\Delta_j \leftarrow -1$. Apply Algorithm 5.4 to part j to get $R[0, j]$, d, and $region\_t[0, \beta_{l_j+1}..\beta_{l_{j+1}-1}]$. $R[0, j] \leftarrow R[0, j] + D^* \cdot |d|$. Determine $region\_t[0, \beta_{l_j}]$ by $\beta_{l_{j+1}}$ and d.

Step 4: Let $\Delta_j \leftarrow 1$. Apply Algorithm 5.4 to part j to get $R[1, j]$, d, and $region\_t[1, \beta_{l_j+1}..\beta_{l_{j+1}-1}]$. $R[1, j] \leftarrow R[1, j] + D^* \cdot |d|$. Determine $region\_t[1, \beta_{l_j}]$ by $\beta_{l_{j+1}}$ and d.

Step 5: If $R[0, j] \leq R[1, j]$, then $D^* \leftarrow R[0, j]$ and $\Delta^*[j] \leftarrow 0$. Otherwise, $D^* \leftarrow R[1, j]$ and $\Delta^*[j] \leftarrow 1$.

Step 6: $j \leftarrow j - 1$. If $j \geq 0$, go to Step 3.

Step 7: $j \leftarrow 0$.

Step 8: $region[\beta_{l_j+1}..\beta_{l_{j+1}-1}] \leftarrow region\_t[\Delta[j], \beta_{l_j+1}..\beta_{l_{j+1}-1}]$.

Step 9: $j \leftarrow j + 1$. If $j \leq N$. Go to Step 8.

Return D^*, $region[1..k]$ and Δ^*.

In Steps 3 and 4, the region of $\left(\beta_{l_{j+1}-1}, \beta_{l_{j+1}}\right)$ has two possible values since $\beta_{l_{j+1}} = \pm\frac{5}{3}$. The region of $\left(\beta_{l_{j+1}-1}, \beta_{l_{j+1}}\right)$ is determined by $\beta_{l_{j+1}}$ and d. If $\beta_{l_{j+1}} = \frac{5}{3}$, then $\left(\beta_{l_{j+1}-1}, \beta_{l_{j+1}}\right)$ is in region 6 if $d > 0$ and in region 5 otherwise. When $\beta_{l_{j+1}} = -\frac{5}{3}$, then $\left(\beta_{l_{j+1}-1}, \beta_{l_{j+1}}\right)$ is in region 2 if $d > 0$ and in region 3 otherwise.

Algorithm for solving problem (DE)

Assume $\boldsymbol{\beta}^*$ is an optimal solution for the problem (DE). Once it is known which region each $\left(\beta_{i-1}^*, \beta_i^*\right)$ is in, $i = 1, \ldots, k$, the original problem (DE) is reduced to a one-dimensional nonlinear optimization problem. There are many one-dimensional search (called linear search) techniques [10] for solving such a problem. Let $region[1..k]$ indicate the region of $\boldsymbol{\beta}^*$. Due to the character of the problem (DE), we assume that there is a given interval $[a, b]$ such that $\beta_0^* \in [a, b]$. Let the solution determined by $\beta_0 = a$ and $region[1..k]$ be $\boldsymbol{\beta}^{(a)}$, and let the solution determined by $\beta_0 = b$ and $region[1..k]$ be $\boldsymbol{\beta}^{(b)}$. If the interval $[a, b]$ is appropriately set, both $\boldsymbol{\beta}^{(a)}$ and $\boldsymbol{\beta}^{(b)}$ are feasible solutions. Hence, $\Phi\left(\beta_0\right)$ is a one-dimensional function with continuous derivative defined over $[a, b]$. Furthermore, $\beta_0^* \in [a, b]$ implies $\Phi'(a) < 0$, $\Phi'(b) > 0$, and $\Phi'\left(\beta_0^*\right) = 0$. Therefore, minimizing the problem (DE) is reduced to solving

$$\Phi'(c) = 0, \quad c \in [a, b] \ , \tag{5.38}$$

that is, finding the root of a nonlinear function. One algorithm for this kind of root-finding problem is

Algorithm 5.6 (Find a root of a nonlinear function)

Given $[a, b]$ such that $\Phi'(a) < 0$ and $\Phi'(b) > 0$.

Step 1: Get a point $c \in (a, b)$.

Step 2: Calculate $\Phi'(c)$.

Step 3: If $|\Phi'(c)| < \epsilon_1$ or $b - a < \epsilon_2$, stop.

Step 4: If $\Phi'(c) > 0$ then let $b \leftarrow c$. Otherwise, let $a \leftarrow c$. Go to Step 1.

We propose two methods to obtain point c in Step 1. The first one is linear interpolation, that is, to approximate $\Phi'\left(\beta_0\right)$ by a linear function passing through $(a, \Phi'(a))$ and $(b, \Phi'(b))$, and let c be the root

$$c = \frac{a\Phi'(b) - b\Phi'(a)}{\Phi'(b) - \Phi'(a)}. \tag{5.39}$$

of this linear function. The second approach is cubic interpolation. It is to approximate $\Phi\left(\beta_0\right)$ by a one-dimensional cubic function determined by triples $(a, \Phi(a), \Phi'(a))$ and $(b, \Phi(b), \Phi'(b))$ and to minimize this cubic function. Numerical experiments indicate that these two approaches require almost the same number of iterations before stopping. However, the cubic interpolation approach needs to calculate the function values $\Phi(a)$ and $\Phi(b)$, which costs more operations. As a consequence, we use formula (5.39) to get point c in Step 1 of Algorithm 5.6.

The next task is to find the interval $[a, b]$ and to determine the region of each $\left(\beta_{i-1}^*, \beta_i^*\right)$, $i = 1, \ldots, k$, i.e., to determine the array $region[1..k]$ of $\boldsymbol{\beta}^*$. These two issues are closely related to each other.

On the one hand, once a specified value of $region[1..k]$ is given, the maximum interval $[a, b]$ of β_0 is determined by $region[1]$ through the definition of the different regions (5.19)-(5.22). They are summarized in Table $T\_interval[0..1, 1..6]$. The maximum value of interval $[a, b]$ is

$$a = T\_interval\,[0, region[1]]\,, \quad b = T\_interval\,[1, region[1]]\,.$$

On the other hand, suppose that the interval $[a, b]$ and $region[1..k]$ are given, and, therefore, that $\boldsymbol{\beta}^{(a)}$ and $\boldsymbol{\beta}^{(b)}$ can be obtained. For $\boldsymbol{\beta}^{(a)}$, we can apply Algorithm 5.5 to obtain its corresponding $region^{(a)}[1..k]$, $D^{*(a)}$ and $\Delta^{*(a)}$. If $D^{*(a)} \geq 0$, then $\boldsymbol{\beta}^{(a)}$ is a local minimizer. Otherwise, a better solution can be obtained if we consider $\boldsymbol{\beta}^{(a)}$ as a feasible solution with $region^{(a)}[1..k]$. That is to say, the current $region[1..k]$ should be updated to $region^{(a)}[1..k]$ in order to reduce the objective function value. Hence, the current $region[1..k]$ could induce the optimal solution only if $region^{(a)}[1..k] = region[1..k]$ and $region^{(b)}[1..k] = region[1..k]$, which implies $region^{(a)}[1..k] = region^{(b)}[1..k]$.

If $region^{(a)}[1..k] \neq region^{(b)}[1..k]$, we get a new value of array $region[1..k]$ and then a new interval $[a, b]$. Checking the solutions corresponding to the endpoints a and b again indicates whether the current $region[1..k]$ gives the right region of $\boldsymbol{\beta}^*$. This process stops when we finally obtain the right interval $[a, b]$ as required in (5.38). The process must stop in a finite number of iterations since the array $region[1..k]$ may have no more than 6^{k+1} possible values.

Table 5.2: $T\_interval[0..1, 1..6]$: the maximum feasible interval of β_0

region of (β_0, β_1)	1	2	3	4	5	6
a	1	1	-1	$-\frac{5}{3}$	$-\frac{5}{3}$	-1
b	$\frac{5}{3}$	$\frac{5}{3}$	1	-1	-1	1

The validation of segmentation must be considered also. For convenience of notation, let the variable previous to β_0 be β_{-1}. By the definition of segmentation, (β_{-1}, β_0) must be an interior point of Ω. Then from the value of β_{-1}, we can calculate the range of β_0 to guarantee that (β_{-1}, β_0) is an interior point of Ω. Let that range be $[l_0, u_0]$. The possible value of β_0 can only be

$$\beta_0 \in [a, b] \cap [l_0, u_0] .$$

Assume $a < l_0$. If, at the feasible solution $\boldsymbol{\beta}$ related to $\beta_0 = l_0$, we get $\Delta^*[0] = -1$ after applying Algorithm 5.5, then β_0 should be decreased to get a better solution. However, $\beta_0 = l_0$ can not be further decreased without changing the value of β_{-1}. Hence, the current segment should be merged with the previous one. Similarly, a bound $\beta_k \in [l_k, u_k]$ is determined by the first variable β_{k+1} of the following segment.

We have the following complete algorithm for solving problem (DE)

Algorithm 5.7 (Complete algorithm for solving (DE))

Let $\boldsymbol{\beta}$ and $region[1..k]$ be given.

Step 1: $\boldsymbol{\beta}^{(a)} \leftarrow \boldsymbol{\beta}$ and $code \leftarrow 0$. Apply Algorithm 5.5 to get $region^{(a)}[1..k]$, $D^{*(a)}$, and $\Delta^{*(a)}$. If $D^{*(a)} \geq 0$, $code \leftarrow 1$ and go to Step 12. Otherwise, $\hat{a} \leftarrow \beta_k^{(a)}$, calculate l_0, u_0, l_k, u_k, and continue.

Step 2: $a \leftarrow T\_interval\left[0, region^{(a)}[1]\right]$ and $b \leftarrow T\_interval\left[1, region^{(a)}[1]\right]$. If $\Delta^{*(a)}[0] > 0$ then $a \leftarrow \beta_0^{(a)}$, otherwise $b \leftarrow \beta_0^{(a)}$. If $\beta_0^{(a)} == b$, then $a \rightleftharpoons b$. $b \leftarrow \min\{b, u_0\}$. $b \leftarrow \max\{b, l_0\}$.

Step 3: Calculate $\boldsymbol{\beta}^{(b)}$ using b and $region^{(a)}[1..k]$.

Step 4: Apply Algorithm 5.5 to get $region^{(b)}[1..k]$, $D^{*(b)}$, and $\Delta^{*(b)}$. Let $\hat{b} \leftarrow \beta_k^{(b)}$.

Step 5.1: If $l_k + \epsilon_1 < \hat{b} < u_k - \epsilon_1$ then go to Step 6.1. Otherwise, continue.

Step 5.2: $\hat{b} \leftarrow \min\left\{\hat{b}, u_k\right\}$. $\hat{b} \leftarrow \max\left\{\hat{b}, l_k\right\}$. Calculate $\boldsymbol{\beta}^{(b)}$ backwards using $\hat{b}$ and $region^{(a)}[1..k]$.

Step 5.3: Apply the backward version of Algorithm 5.5 to get $region^{(b)}[1..k]$, $D^{*(b)}$ and $\Delta^{*(b)}$.

Step 5.4: If $\left(\hat{b} - \hat{a}\right) \cdot \Delta^{*(b)}[0] > 0$, then $code \leftarrow 3$, $\boldsymbol{\beta} \leftarrow \boldsymbol{\beta}^{(b)}$, $region[1..k] \leftarrow region^{(b)}[1..k]$, and go to Step 12. Otherwise, $b \leftarrow \beta_0^{(b)}$ and go to Step 7.

Step 6.1: If $D^{*(b)} \geq 0$, then $code \leftarrow 1$, $\boldsymbol{\beta} \leftarrow \boldsymbol{\beta}^{(b)}$, $region[1..k] \leftarrow region^{(b)}[1..k]$, and go to Step 12.

Step 6.2: If ($b == l_0$ or $b == u_0$) and $(b-a)\cdot\Delta^{*(b)}[0] > 0$, let $code \leftarrow 2$, $\boldsymbol{\beta} \leftarrow \boldsymbol{\beta}^{(b)}$ and $region[1..k] \leftarrow region^{(b)}[1..k]$, go to Step 12. Otherwise, go to Step 7.

Step 6.3: If $(b-a)\cdot\Delta^{*(b)}[0] < 0$ and $region^{(a)}[1..k] == region^{(b)}[1..k]$, go to Step 7. Otherwise, $a \leftarrow b$, $\boldsymbol{\beta}^{(a)} \leftarrow \boldsymbol{\beta}^{(b)}$, $region^{(a)}[1..k] \leftarrow region^{(b)}[1..k]$, $D^{*(a)} \leftarrow D^{*(b)}$, $\Delta^{*(a)} \leftarrow \Delta^{*(b)}$, $\hat{a} \leftarrow \hat{b}$, and go to Step 2.

Step 7: $region[1..k] \leftarrow region^{(b)}[1..k]$. If $a > b$, then $a \rightleftharpoons b$, $\beta^{(a)} \rightleftharpoons \beta^{(b)}$, and $D^{*(a)} \rightleftharpoons D^{*(b)}$. $\Phi'(a) \leftarrow D^{*(a)}$ and $\Phi'(b) \leftarrow -D^{*(b)}$.

Step 8: $c \leftarrow \frac{a\Phi'(b) - b\Phi'(a)}{\Phi'(b) - \Phi'(a)}$. Calculate $\boldsymbol{\beta}^{(c)}$ using c and $region[1..k]$.

Step 9: Apply Algorithm 5.3 to $\boldsymbol{\beta}^{(c)}$ and $region[1..k]$) to get $\Phi'(c)$.

Step 10: If $|\Phi'(c)| < \epsilon_2$ or $b - a < \epsilon_3$, then $code \leftarrow 1$, $\boldsymbol{\beta} \leftarrow \boldsymbol{\beta}^{(c)}$, and go to Step 12.

Step 11: If $\Phi'(c) > 0$ then $b \leftarrow c$, $\boldsymbol{\beta}^{(b)} \leftarrow \boldsymbol{\beta}^{(c)}$ and $\Phi'(b) \leftarrow \Phi'(c)$. Otherwise, $a \leftarrow c$, $\boldsymbol{\beta}^{(a)} \leftarrow \boldsymbol{\beta}^{(c)}$, $\Phi'(a) \leftarrow \Phi'(c)$, and go to Step 8.

Step 12: Return $\boldsymbol{\beta}$, $region[1..k]$ and $code$. Stop.

In Algorithm 5.7, the variable *code* indicates the status of returned variables. If $code = 1$, then the returned $\boldsymbol{\beta}$ is an optimal solution of problem (DE) and the region of each variable pair is indicated by the returned $region[1..k]$. If $code = 2$, then the current segment merges with the previous segment. If $code = 3$, then the current segment merges with the following segment.

Steps $1-6$ of Algorithm 5.7 form the first iteration, which determines the region of $\boldsymbol{\beta}^*$ and the interval $[a, b]$. Step 1 initializes information related to one endpoint a. In the first iteration, the first variable β_0 is always moved from a towards b. The last variable β_k takes on values in $\left[\hat{a}, \hat{b}\right]$ and β_k is always moved from $\hat{a}$ towards $\hat{b}$. Step 5 handles the last variable β_k. Once the endpoint $\hat{b}$ takes on the value of l_k or u_k, the first iteration terminates in Step 5.4. There are two possible cases. The first one is that β_k needs to be moved away from $\hat{a}$, which means that the current segment must merge with the following segment and the algorithm stops here. The second case is that β_k needs to be moved towards $\hat{a}$. Then we obtain the region of $\boldsymbol{\beta}^*$ and the interval including $\boldsymbol{\beta}^*$. The algorithm jumps to the second iteration to actually find $\boldsymbol{\beta}^*$. Step 6 handles the first variable β_0. The same logic as in Step 5 applies here. Steps $7-11$ form the second iteration to calculate $\boldsymbol{\beta}^*$. This iteration is actually a specific version of Algorithm 5.6.

5.3.3 Calculate Lagrange multipliers

In this section, we only consider one segment. For convenience of notation, we assume that the optimal solution of problem (DE) obtained in the previous section is $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)$. By Theorem 5.4, $\boldsymbol{\beta}$ satisfies

$$-b_i = n_2\left(\beta_{i-1}, \beta_i; \lambda_i, \mu_i\right) + n_1\left(\beta_i, \beta_{i+1}; \lambda_{i+1}, \mu_{i+1}\right), \quad i = 0, \ldots, k, \tag{5.40}$$

where the λ_i and μ_i are Lagrange multipliers corresponding to the constraint $g\left(\beta_{i-1}, \beta_i\right) \leq 0$. The Lagrange multipliers λ_i and μ_i can be either positive or negative.

If (5.40) can be satisfied only when one of λ_i and μ_i is negative, then the constraint $g\left(\beta_{i-1}, \beta_i\right) \leq 0$ can not be active at the dual optimal solution. Hence, the current segment should be split at the variable pair $\left(\beta_{i-1}, \beta_i\right)$ to form a new segmentation.

Assume that there is at least one axis point in $\boldsymbol{\beta}$. For example, assume $\beta_l = \frac{5}{3}$. Then we must have $\beta_{l-1} = -1$ and $\beta_{l+1} = -1$. Both $\left(\beta_{l-1}, \beta_l\right)$ and $\left(\beta_l, \beta_{l+1}\right)$ are axis points. The point $\left(\beta_{l-1}, \beta_l\right) = \left(-1, \frac{5}{3}\right)$ is in either region 5 or region 6. Applying formula (5.11), we obtain

$$\begin{pmatrix} n_1\left(\beta_{l-1}, \beta_l; \lambda_l, \mu_l\right) \\ n_2\left(\beta_{l-1}, \beta_l; \lambda_l, \mu_l\right) \end{pmatrix} = \lambda_l \begin{pmatrix} \beta_{l-1} + \beta_l - \frac{2}{3} \\ \beta_{l-1} + \beta_l + \frac{2}{3} \end{pmatrix} = \lambda_l \begin{pmatrix} 0 \\ \frac{4}{3} \end{pmatrix}.$$

By changing the value of λ_l to $\frac{4}{3}\lambda_l$, this formula can be restated as

$$\begin{pmatrix} n_1\left(\beta_{l-1}, \beta_l; \lambda_l, \mu_l\right) \\ n_2\left(\beta_{l-1}, \beta_l; \lambda_l, \mu_l\right) \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda_l \end{pmatrix}.$$

Similarly, applying formula (5.10) to the point $\left(\beta_l, \beta_{l+1}\right) = \left(\frac{5}{3}, -1\right)$, one obtains

$$\begin{pmatrix} n_1\left(\beta_l, \beta_{l+1}; \lambda_{l+1}, \mu_{l+1}\right) \\ n_2\left(\beta_l, \beta_{l+1}; \lambda_{l+1}, \mu_{l+1}\right) \end{pmatrix} = \begin{pmatrix} \lambda_{l+1} \\ 0 \end{pmatrix}.$$

Therefore, the conditions of (5.40) that are related to β_{l-1}, β_l and β_{l+1} are

$$-b_{l-1} = n_2\left(\beta_{l-2}, \beta_{l-1}; \lambda_{l-1}, \mu_{l-1}\right),$$

$$-b_l = \lambda_l + \lambda_{l+1},$$

$$-b_{l+1} = n_1\left(\beta_{l+1}, \beta_{l+2}; \lambda_{l+1}, \mu_{l+1}\right).$$

Clearly the values of λ_l and λ_{l+1} are independent of all other Lagrange multipliers and vice versa. If $b_l \le 0$, then there exist non-negative λ_l and λ_{l+1} that satisfy (5.40). If $b_l > 0$, then, no matter what the value of $\lambda_l \ge 0$ is, we must have $\lambda_{l+1} = -b_l - \lambda_l < 0$. A similar result holds when $\beta_l = -\frac{5}{3}$: There are non-negative λ_l and λ_{l+1} satisfying (5.40) if and only if $b_l \ge 0$. In summary, we have the following theorem.

Theorem 5.10 *If $\beta_l = \pm\frac{5}{3}$, then the constraints $g(\beta_{l-1}, \beta_l) \le 0$ and $g(\beta_l, \beta_{l+1}) \le 0$ are active at the dual optimal solution if and only if*

$$b_l \cdot sign(\beta_l) \le 0. \tag{5.41}$$

In the rest of this subsection, we assume that there are no axis points in $\boldsymbol{\beta}$.

Assume that each (β_{i-1}, β_i) is neither a corner point nor an axis point. The normal cone of each variable pair (β_{i-1}, β_i), $i = 1, \ldots, k$, is defined by (5.10) or (5.11). Both (5.10) and (5.11) yield

$$\begin{pmatrix} n_1(\beta_{i-1}, \beta_i; \lambda_i, \mu_i) \\ n_2(\beta_{i-1}, \beta_i; \lambda_i, \mu_i) \end{pmatrix} = \lambda_i \begin{pmatrix} n_1(\beta_{i-1}, \beta_i; 1, 0) \\ n_2(\beta_{i-1}, \beta_i; 1, 0) \end{pmatrix}.$$

Consequently, (5.40) defines a linear system of $k+1$ equations and k unknown variables $\lambda_1, \ldots, \lambda_k$. The coefficient matrix of this linear system is

$$\begin{bmatrix} n_1(\beta_0, \beta_1; 1, 0) & & & & \\ n_2(\beta_0, \beta_1; 1, 0) & n_1(\beta_1, \beta_2; 1, 0) & & & \\ & n_2(\beta_1, \beta_2; 1, 0) & n_1(\beta_2, \beta_3; 1, 0) & & \\ & & \ddots & \ddots & \\ & & & n_2(\beta_{k-2}, \beta_{k-1}; 1, 0) & n_1(\beta_{k-1}, \beta_k; 1, 0) \\ & & & & n_2(\beta_{k-1}, \beta_k; 1, 0) \end{bmatrix}_{k\times(k+1)}. \tag{5.42}$$

(β_{i-1}, β_i) not being an axis point ensures that both $n_1(\beta_{i-1}, \beta_i; 1, 0)$ and $n_2(\beta_{i-1}, \beta_i; 1, 0)$ are not zero. Such a linear system is easily solved by the following algorithm

Algorithm 5.8 (Calculation of Lagrange multipliers when there are no corner points or axis points)

Let $\boldsymbol{\beta}$, $region[1..k]$ and $B = (b_0, \ldots, b_k)$ be given.

Step 0: $r \leftarrow 0$, $i \leftarrow 1$.

Step 1: $u \leftarrow n_1(\beta_{i-1}, \beta_i; 1, 0)$, and $v \leftarrow n_2(\beta_{i-1}, \beta_i; 1, 0)$.

Step 2: $\lambda[i] \leftarrow (-b[i-1] - r)/u$.

Step 3: $r \leftarrow \lambda[i] \cdot v$.

Step 4: $i \leftarrow i + 1$. If $i \leq k$, go to Step 1.

Return $\lambda[1..k]$.

Next we consider the situation when at least one (β_{i-1}, β_i) is a corner point and there are no axis points. By Theorem 5.6, this situation will occur when all of the β_i are 1 or -1. Without loss of generality, assume that $\beta_i = 1$, for all $i = 0, \ldots, k$. From (5.12),

$$\begin{pmatrix} n_1(\beta_{i-1}, \beta_i; \lambda, \mu) \\ n_2(\beta_{i-1}, \beta_i; \lambda, \mu) \end{pmatrix} = \lambda_i \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu_i \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \lambda_i \\ \mu_i \end{pmatrix}. \tag{5.43}$$

Hence, a normal vector of each (β_{i-1}, β_i) is determined by two Lagrange multipliers λ_i and μ_i. Condition (5.40) defines a linear system with $k+1$ equations, $2k$ unknown Lagrange multipliers and $k-1$ degrees of freedom.

If we consider the normal vectors at (β_{i-1}, β_i) as variables and denote them by $n_1(\beta_{i-1}, \beta_i)$ and $n_2(\beta_{i-1}, \beta_i)$, then formula (5.43) defines a one-to-one onto transformation between $n_1(\beta_{i-1}, \beta_i)$, $n_2(\beta_{i-1}, \beta_i)$ and λ_i, μ_i. The reverse transformation is

$$\begin{pmatrix} \lambda_i \\ \mu_i \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} n_1(\beta_{i-1}, \beta_i) \\ n_2(\beta_{i-1}, \beta_i) \end{pmatrix}. \tag{5.44}$$

From (5.43) and (5.44), we can show that $\lambda_i \geq 0$ and $\mu_i \geq 0$ if and only if

$$\begin{gathered} n_1(\beta_{i-1}, \beta_i) \geq 0, \\ n_2(\beta_{i-1}, \beta_i) \geq 0, \\ \tfrac{1}{2} n_1(\beta_{i-1}, \beta_i) \leq n_2(\beta_{i-1}, \beta_i) \leq 2 n_1(\beta_{i-1}, \beta_i). \end{gathered} \tag{5.45}$$

The set defined by (5.45) is the shaded area in Figure 5.4. Hence, checking the sign of the Lagrange multipliers is equivalent to checking if condition (5.45) holds. If there exist $n_1(\beta_{i-1}, \beta_i)$ and $n_2(\beta_{i-1}, \beta_i)$

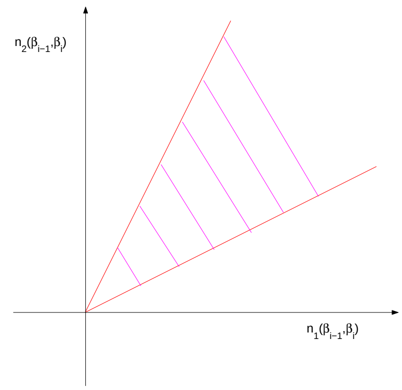

Figure 5.4: Feasible set of normal vectors when all $\beta_i = 1$.

such that all conditions (5.40) and (5.45) are satisfied, then $\boldsymbol{\beta}$ is an optimal solution of the current segment and, therefore, $\boldsymbol{\beta}$ is a part of the dual optimal solution. Otherwise, at least one of the Lagrange multipliers is negative and the current segment should be split.

When $i = 0$, condition (5.40) becomes

$$-b_0 = n_1\left(\beta_0, \beta_1\right). \tag{5.46}$$

Hence, the value of $n_1\left(\beta_0, \beta_1\right)$ is uniquely determined. If $b_0 > 0$, then the constraint $g\left(\beta_0, \beta_1\right) \leq 0$ can not be active at the dual optimal solution and we need to split the current segment at $\left(\beta_0, \beta_1\right)$ to get a new segmentation.

For given $n_1\left(\beta_0, \beta_1\right) \geq 0$, the value of $n_2\left(\beta_0, \beta_1\right)$ satisfying condition (5.45) is in the interval

$$\left[\frac{1}{2}n_1\left(\beta_0, \beta_1\right), 2n_1\left(\beta_0, \beta_1\right)\right] = \left[-\frac{1}{2}b_0, -2b_0\right]. \tag{5.47}$$

Furthermore, the value of $n_2\left(\beta_0, \beta_1\right)$ should satisfy condition (5.40). When $i = 1$, condition (5.40) becomes

$$-b_1 = n_2\left(\beta_0, \beta_1\right) + n_1\left(\beta_1, \beta_2\right). \tag{5.48}$$

If $n_1\left(\beta_1, \beta_2\right) \geq 0$, then (5.48) implies

$$n_2\left(\beta_0, \beta_1\right) \in \left[0, -b_1\right]. \tag{5.49}$$

Hence, any feasible $n_2\left(\beta_0, \beta_1\right)$ must be inside the intersection of the two intervals in (5.47) and (5.49). If $-b_1 < -\frac{1}{2}b_0$, then there does not exist an $n_1\left(\beta_1, \beta_2\right)$ that satisfies both (5.40) and (5.45) and we need to split the current segment at (β_1, β_2) to get a better dual solution. If $-b_1 \geq -\frac{1}{2}b_0$, then the value of $n_2(\beta_0, \beta_1)$ is inside the interval

$$\left[-\frac{1}{2}b_0, -2b_0\right] \cap [0, -b_1] = \left[-\frac{1}{2}b_0, \min\{-2b_0, -b_1\}\right].$$

Let $[l_i, u_i]$ denote the feasible interval of $n_2(\beta_{i-1}, \beta_i)$, $i = 1, \ldots, k-1$. We have just shown that

$$[l_1, u_1] = \left[-\frac{1}{2}b_0, \min\{-2b_0, -b_1\}\right]. \tag{5.50}$$

In general, once the feasible interval $[l_{i-1}, u_{i-1}]$ of $n_2(\beta_{i-2}, \beta_{i-1})$ is known, we are able to get the feasible interval of $n_1(\beta_{i-1}, \beta_i)$ to be $[-b_{i-1} - u_{i-1}, -b_{i-1} - l_{i-1}]$ by condition (5.40), i.e.,

$$-b_{i-1} = n_2(\beta_{i-2}, \beta_{i-1}) + n_1(\beta_{i-1}, \beta_i).$$

Further applying condition (5.45), the feasible interval of $n_2(\beta_{i-1}, \beta_i)$ is obtained to be

$$[l_i, u_i] = \left[-\frac{1}{2}b_{i-1} - \frac{1}{2}u_{i-1}, \min\{-2b_{i-1} - 2l_{i-1}, -b_i\}\right].$$

If $-b_i < l_i$, then the current segment should be split at (β_i, β_{i+1}). This process is continued until all $[l_i, u_i]$, $i = 1, \ldots, k-1$, are obtained.

If we perform the process one more step, we find that the feasible interval of $n_2(\beta_{k-1}, \beta_k)$ is

$$n_2(\beta_{k-1}, \beta_k) \in \left[-\frac{1}{2}b_{k-1} - \frac{1}{2}u_{k-1}, -2b_{k-1} - 2l_{k-1}\right]. \tag{5.51}$$

Then we consider the condition (5.40) when $i = k+1$, which yields

$$-b_k = n_2(\beta_{k-1}, \beta_k). \tag{5.52}$$

However, in general $\left[-\frac{1}{2}b_{k-1} - \frac{1}{2}u_{k-1}, -2b_{k-1} - 2l_{k-1}\right]$ is a real interval containing more than one point. That means the interval $[l_i, u_i]$ for each $n_2(\beta_{i-1}, \beta_i)$ is not tight. We need to run the above procedure backwards from k to 1.

In this backward process, after $n_2(\beta_i, \beta_{i+1}) \in [l_{i+1}, u_{i+1}]$ is obtained, the feasible interval of $n_1(\beta_i, \beta_{i+1})$ is, by condition (5.45), $\left[\frac{1}{2}l_{i+1}, 2u_{i+1}\right]$. Therefore, by condition (5.40), a new interval for $n_2(\beta_{i-1}, \beta_i)$ is

$$n_2(\beta_{i-1}, \beta_i) \in \left[\max\{-b_i - 2u_{i+1}, 0\}, -b_i - \frac{1}{2}l_{i+1}\right].$$

Now take the intersection of this interval with the interval obtained by the forward process to obtain a tight feasible interval for $n_2(\beta_{i-1}, \beta_i)$, namely,

$$[l_i, u_i] \leftarrow [l_i, u_i] \cap \left[\max\{-b_i - 2u_{i+1}, 0\}, -b_i - \frac{1}{2}l_{i+1}\right].$$

Since $l_i \geq 0$ is guaranteed by the forward process, the above operation can be simplified:

$$[l_i, u_i] \leftarrow [l_i, u_i] \cap \left[-b_i - 2u_{i+1}, -b_i - \frac{1}{2}l_{i+1}\right].$$

Furthermore, the condition

$$-b_i = n_2(\beta_{i-1}, \beta_i) + n_1(\beta_i, \beta_{i+1})$$

implies

$$n_2(\beta_{i-1}, \beta_i) = -b_i - n_1(\beta_i, \beta_{i+1}).$$

Hence, if $-b_i < \min n_1(\beta_i, \beta_{i+1}) = \frac{1}{2}l_{i+1}$, then there does not exist a non-negative $n_2(\beta_{i-1}, \beta_i)$ satisfying condition (5.40). The current segment should be split at the variable pair (β_{i-1}, β_i).

In summary, we have the following algorithm to check if the current segment should be split. If not, the feasible value of $n_2(\beta_{i-1}, \beta_i)$ is in the interval $[l_i, u_i]$.

Algorithm 5.9 (Calculate Lagrange multipliers when all $\beta_i = 1$)

Let $b[0..k]$ be given.

Step 0: $bSplit \leftarrow false$, $I \leftarrow 0$.

Step 1: If $b[0] > 0$ then $bSplit \leftarrow true$, $I \leftarrow 1$ and go to Step 7. Otherwise, $l[0] \leftarrow 0$ and $u[0] \leftarrow 0$.

For $i = 1..k-1$

Step 2: $l[i] \leftarrow -\frac{1}{2} \cdot (b[i-1] + u[i-1])$ and $u[i] \leftarrow \min\{-2 \cdot (b[i-1] + l[i-1]), -b[i]\}$.

Step 3: If $-b[i] < l[i]$ then $bSplit \leftarrow true$, $I \leftarrow i+1$ and go to Step 7.

end For

Step 4: If $b[k] > 0$ then $bSplit \leftarrow true$, $I \leftarrow k$ and go to Step 7. Otherwise, $l[k] \leftarrow -b[k]$ and $u[k] \leftarrow -b[k]$.

For $i = k-1..1$

Step 5: $l[i] \leftarrow \max\{l[i], -b[i] - 2u[i+1]\}$ and $u[i] \leftarrow \min\left\{u[i], -b[i] - \frac{1}{2}l[i+1]\right\}$.

Step 6: If $-b[i] < \frac{1}{2}l[i+1]$ then $bSplit \leftarrow true$, $I \leftarrow i$ and go to Step 7.

end For

Step 7: Return $bSplit$ and I. Stop.

At the end of the algorithm, if the boolean variable $bSplit$ is $true$, then we need to split the current segment at the variable pair (β_{I-1}, β_I).

Example 5.11 *For a segment with 3 variables, assume that the coefficients of the objective function are* $b_0 = -1$*,* $b_1 = -2$*,* $b_2 = -1$*. The solution obtained from the previous section is* $\beta_0 = \beta_1 = \beta_2 = 1$*.*

For this example, the linear system defined by (5.40) is

$$1 = n_1(\beta_0, \beta_1),$$

$$2 = n_2(\beta_0, \beta_1) + n_1(\beta_1, \beta_2),$$

$$1 = n_2(\beta_1, \beta_2).$$

Since $n_1(\beta_1, \beta_2)$ *can be expressed as*

$$n_1(\beta_1, \beta_2) = 2 - n_2(\beta_0, \beta_1)\ ,$$

there is one free variable $n_2(\beta_0, \beta_1)$*. Condition (5.45) is*

$$\frac{1}{2} \leq n_2(\beta_0, \beta_1) \leq 2,$$

$$0 \leq 2 - n_2(\beta_0, \beta_1),$$

$$1 - \frac{1}{2}n_2(\beta_0, \beta_1) \leq 1 \leq 4 - 2n_2(\beta_0, \beta_1)\ ,$$

which after simplification yields

$$\frac{1}{2} \leq n_2(\beta_0, \beta_1) \leq \frac{3}{2}.$$

Therefore, $\beta = (1, 1, 1)^T$ *is the dual optimal solution. The free variable* $n_2(\beta_0, \beta_1)$ *can be any number in* $\left[\frac{1}{2}, \frac{2}{3}\right]$*.*

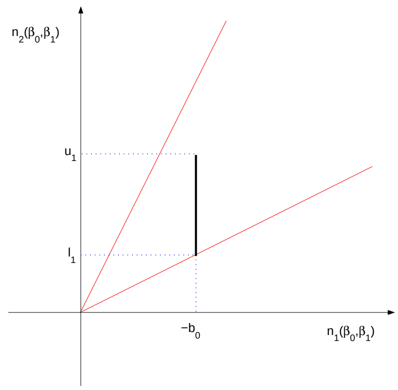

Figure 5.5: Feasible set E_1.

Now we apply Algorithm 5.9. First consider the forward process, which consists of steps $1-3$*. For* $i = 1$*,* $b_0 = -1 > 0$ *is not true, the interval* $[l_1, u_1]$ *is given by*

$$l_1 = -\frac{1}{2}b_0 = \frac{1}{2},$$

$$u_1 = \min\{-b_1, -2b_0\} = 2,$$

and $-b_1 = 2 < l_1 = \frac{1}{2}$ *is not true.*

Next consider the backward process, which consists of the steps $4-6$*. For* $i = 1$*,* $b_2 = -1 > 0$ *is not satisfied. The new interval for* $n_2(\beta_0, \beta_1)$ *is*

$$\left[-b_1 - 2(-b_2), -b_1 - \frac{1}{2}(-b_2)\right] = \left[0, \frac{3}{2}\right].$$

Therefore, the tight feasible interval for $n_2(\beta_0, \beta_1)$ *is*

$$[l_1, u_1] = \left[\frac{1}{2}, 2\right] \cap \left[0, \frac{3}{2}\right] = \left[\frac{1}{2}, \frac{3}{2}\right].$$

The condition $-b_1 = 2 < \frac{1}{2}l_2 = -\frac{1}{2}b_2 = \frac{1}{2}$ *is not satisfied. Algorithm 5.9 gives the same result as that obtained by solving for the Lagrange multipliers directly.*

If the value of *bSplit* returned by Algorithm 5.9 is *true*, i.e., if there exist $n_1(\beta_{i-1}, \beta_i)$ and $n_2(\beta_{i-1}, \beta_i)$, $i = 1, \ldots, k$, that satisfy all conditions (5.40) and (5.45), then the feasible set for $n_1(\beta_{i-1}, \beta_i)$ and $n_2(\beta_{i-1}, \beta_i)$ can be expressed explicitly as follows. A vector $(n_1(\beta_0, \beta_1), n_2(\beta_0, \beta_1))^T$ satisfies con-

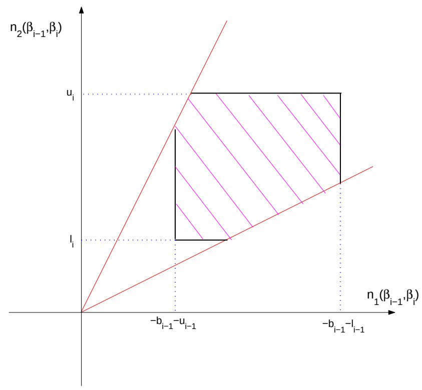

Figure 5.6: Feasible set $E_i, i = 2, \ldots, k$.

ditions (5.40) and (5.45) if and only if it is in the set

$$E_1 = \{\ n_1\left(\beta_0, \beta_1\right) = -b_0,\ n_2\left(\beta_0, \beta_1\right) \in \left[l_1, u_1\right]\ \}\,. \tag{5.53}$$

This set E_1 is the solid vertical black line segment in Figure 5.5.

For $i = 2, \ldots k-1$, a vector $\left(n_1\left(\beta_{i-1}, \beta_i\right), n_2\left(\beta_{i-1}, \beta_i\right)\right)^T$ satisfies conditions (5.40) and (5.45) if and only if it is in the set

$$\begin{aligned} E_i = \{\ & n_1\left(\beta_{i-1}, \beta_i\right) \in \left[-b_{i-1} - u_{i-1}, -b_{i-1} - l_{i-1}\right], \\ & n_2\left(\beta_{i-1}, \beta_i\right) \in \left[l_i, u_i\right], \\ & \tfrac{1}{2} n_1\left(\beta_{i-1}, \beta_i\right) \leq n_2\left(\beta_{i-1}, \beta_i\right) \leq 2 n_1\left(\beta_{i-1}, \beta_i\right)\ \}\,. \end{aligned} \tag{5.54}$$

This set is the shaded area in Figure 5.6.

A vector $\left(n_1\left(\beta_{k-1}, \beta_k\right), n_2\left(\beta_{k-1}, \beta_k\right)\right)^T$ satisfies conditions (5.40) and (5.45) if and only if it is in the set

$$E_k = \{\ n_1\left(\beta_{k-1}, \beta_k\right) \in \left[-b_{k-1} - u_{k-1}, -b_{k-1} - l_{k-1}\right],\ n_2\left(\beta_{k-1}, \beta_k\right) = -b_{k+1}\ \}\,. \tag{5.55}$$

The requirements $n_1\left(\beta_{i-1}, \beta_i\right) \geq 0$ and $n_2\left(\beta_{i-1}, \beta_i\right) \geq 0$ are not explicit in the definition of E_i. They are incorporated in the calculation of $[l_i, u_i]$.

When $\beta_i = -1$, for all $i = 0, \ldots, k$, the following similar algorithm can be used.

Algorithm 5.10 (Calculate Lagrange multipliers when all $\beta_i = -1$)

Let $b[0..k]$ be given.

Step 0: $bSplit \leftarrow false$, $I \leftarrow 0$.

Step 1: If $b[0] < 0$ then $bSplit \leftarrow true$, $I \leftarrow 1$ and go to Step 7. Otherwise, $l[0] \leftarrow 0$ and $u[0] \leftarrow 0$.

For $i = 1..k-1$

Step 2: $l[i] \leftarrow \max\{-2 \cdot (b[i-1] + u[i-1]), -b[i]\}$ and $u[i] \leftarrow -\frac{1}{2} \cdot (b[i-1] + l[i-1])$.

Step 3: If $-b[i] > u[i]$ then $bSplit \leftarrow true$, $I \leftarrow i+1$ and go to Step 7.

end For

Step 4: If $b[k] < 0$ then $bSplit \leftarrow true$, $I \leftarrow k$ and go to Step 7. Otherwise, $l[k] \leftarrow -b[k]$ and $u[k] \leftarrow -b[k]$.

For $i = k-1..1$

Step 5: $l[i] \leftarrow \max\left\{l[i], -b[i] - \frac{1}{2}u[i+1]\right\}$ and $u[i] \leftarrow \min\{u[i], -b[i] - 2l[i+1]\}$.

Step 6: If $-b[i] > \frac{1}{2}u[i+1]$, then $bSplit \leftarrow true$, $I \leftarrow i$ and go to Step 7.

end For

Step 7: Return $bSplit$ and I. Stop.

5.3.4 Initialization

Let the initial dual solution $\beta = (\beta_0, \ldots, \beta_{n-2})$ be given by

$$\beta_i = \begin{cases} 1, & \text{if } b_i < 0, \\ -1, & \text{if } b_i \geq 0. \end{cases}$$

It is easy to show that such a dual solution is feasible. However, this solution may not create a valid segmentation, since there may be more than two consecutive dual variable pairs that are interior points of Ω.

The points $(1,1)$ and $(-1,-1)$ are corner points of Ω. Hence, (β_{i-1}, β_i) is an interior point only when $(\beta_{i-1}, \beta_i) = (1,-1)$ or $(-1,1)$. Without loss of generality, assume that $(\beta_{i-1}, \beta_i) = (1,-1)$. Also assume that the pair immediately preceding (β_{i-1}, β_i) is part of a valid segment. That happens only when $i = 0$ or $\beta_{i-2} = 1$.

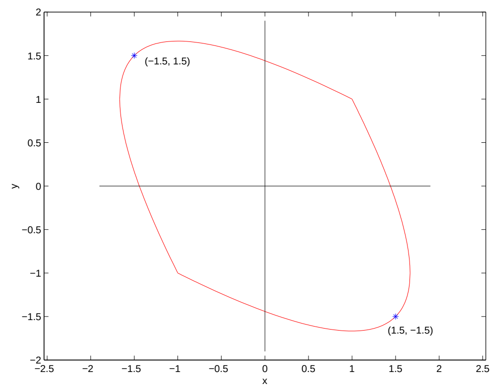

Figure 5.7: Both $\left(\frac{3}{2}, -\frac{3}{2}\right)$ and $\left(-\frac{3}{2}, \frac{3}{2}\right)$ are boundary points of Ω.

If $i = n-2$ or $\beta_{i+1} = -1$, then both $(\beta_{i-2}, \beta_{i-1}) = (1, 1)$ and $(\beta_i, \beta_{i+1}) = (-1, -1)$ are corner points of Ω. Therefore, $(\beta_{i-1}, \beta_i) = (1, -1)$ results in a valid segmentation. In this case, nothing need be done.

If $\beta_{i+1} = 1$, then there are at least two consecutive pairs (β_{i-1}, β_i) and (β_i, β_{i+1}) that are interior points of Ω. Assume that β_{i-1}, β_i, ..., β_{i+k} is a subsequence such that each consecutive pair of this subsequence is an interior point of Ω and that both $(\beta_{i-2}, \beta_{i-1})$ and $(\beta_{i+k}, \beta_{i+k+1})$ are boundary points of Ω. This subsequence must be 1 and -1 alternatively. There are two different cases of such a subsequence.

The first case is $k = 1$. This results in the four consecutive pairs

$$(\beta_{i-2}, \beta_{i-1}) = (1, 1)\,,$$

$$(\beta_{i-1}, \beta_i) = (1, -1)\,,$$

$$(\beta_i, \beta_{i+1}) = (-1, 1)\,,$$

$$(\beta_{i+1}, \beta_{i+2}) = (-1, -1)\,.$$

In this case, it is easy to make the segmentation valid by assigning

$$\beta_i = -\frac{5}{3}.$$

The second case is $k \geq 2$. In this case, we can make the segmentation valid by multiplying every variable in the subsequence by $\frac{3}{2}$, i.e.,

$$\beta_{i+l} \leftarrow \beta_{i+l} \cdot \frac{3}{2}, \quad l = 0, \ldots, k-1.$$

It is easy to prove that both $\left(\frac{3}{2}, -\frac{3}{2}\right)$ and $\left(-\frac{3}{2}, \frac{3}{2}\right)$ are boundary points of Ω by (5.20) and (5.22). These two points are indicated by '*' in Figure 5.7.

The region of each pair (β_{i-1}, β_i) is determined at the same time the value of each β_i is determined.

5.4 Algorithm to obtain primal optimal solution

In this section, we introduce an algorithm for solving the dual-to-primal problem (T) to obtain the primal optimal solution. This algorithm is closely related to the method of calculating the Lagrange multipliers discussed in Section 5.3.3.

It is sufficient to develop the algorithm for one segment only. Let the dual optimal solution corresponding to the current segment be $\boldsymbol{\beta}^* = (\beta_0^*, \ldots, \beta_k^*)$. We need to solve the following dual-to-primal transformation subproblem

$$(\mathrm{TS}) \quad \left\{ \begin{array}{lrcl} \displaystyle\min_{q_i, \lambda_i, \mu_i} & \displaystyle\sum_{i=1}^{k+1} |q_i| & & \\ \text{s.t.} & q_1 & = & \Delta z_0, \\ & q_i & = & \Delta z_i - n_1\left(\beta_{i-1}^*, \beta_i^*; \lambda_i, \mu_i\right), \ i = 1, \ldots, k, \\ & q_{i+1} & = & \Delta z_i + n_2\left(\beta_{i-1}^*, \beta_i^*; \lambda_i, \mu_i\right), \ i = 1, \ldots, k, \\ & q_{k+1} & = & \Delta z_{k+1}, \\ & \lambda_i, \mu_i & \geq & 0, \ i = 1, \ldots, k. \end{array} \right.$$

Assume there is at least one axis point in $\boldsymbol{\beta}^*$. Without loss of generality, assume $\beta_l^* = \frac{5}{3}$. Then we must have $\beta_{l-1}^* = -1$ and $\beta_{l+1}^* = -1$. Both $\left(\beta_{l-1}^*, \beta_l^*\right)$ and $\left(\beta_l^*, \beta_{l+1}^*\right)$ are axis points. The normal vectors at $\left(\beta_{l-1}^*, \beta_l^*\right)$ and $\left(\beta_l^*, \beta_{l+1}^*\right)$ are

$$n_1\left(\beta_{l-1}^*, \beta_l^*; \lambda_i, \mu_i\right) = 0,$$

$$n_2\left(\beta_{l-1}^*, \beta_l^*; \lambda_i, \mu_i\right) \geq 0,$$

$$n_1\left(\beta_l^*, \beta_{l+1}^*; \lambda_i, \mu_i\right) \geq 0,$$

$$n_2\left(\beta_l^*, \beta_{l+1}^*; \lambda_i, \mu_i\right) = 0.$$

As a consequence, by the model (TS), we know that

$$q_l = \Delta z_l, \tag{5.56}$$

$$q_{l+1} = \Delta z_l + n_2\left(\beta_{l-1}^*, \beta_l^*; \lambda_i, \mu_i\right) = \Delta z_{l+1} - n_1\left(\beta_l^*, \beta_{l+1}^*; \lambda_i, \mu_i\right), \tag{5.57}$$

and

$$q_{l+2} = \Delta z_{l+1}. \tag{5.58}$$

Formula (5.56) implies that there is a small-size problem (TS) with variables $q_1, \ldots, q_l$. Its solution is independent of the other q_i. Formula (5.58) implies that we can solve another small-size problem (TS) with variables $q_{l+2}, \ldots, q_{k+1}$. From (5.57), the value of q_{l+1} can be any value between Δz_l and Δz_{l+1}. Furthermore, the value of q_{l+1} is independent of all of the other q_i. Therefore, the value of q_{l+1} is determined by

$$\begin{cases} \min & |q_{l+1}| \\ \text{s.t.} & \Delta z_l \leq q_{l+1} \leq \Delta z_{l+1}. \end{cases}$$

If $\beta_l^* = \pm\frac{5}{3}$, then the value of q_{l+1} is determined by the one-dimensional optimization problem

$$\begin{cases} \min & |q_{l+1}| \\ \text{s.t.} & \min\left\{\Delta z_l, \Delta z_{l+1}\right\} \leq q_{l+1} \leq \max\left\{\Delta z_l, \Delta z_{l+1}\right\} \end{cases},$$

the solution of which is $q_{l+1} = \min\left\{\left|\Delta z_l\right|, \left|\Delta z_{l+1}\right|\right\}$, if Δz_l and Δz_{l+1} have the same sign. Otherwise, $q_{l+1} = 0$.

In the rest of this section, we assume that none of the $\left(\beta_{i-1}^*, \beta_i^*\right)$, $i = 1, \ldots, k$ is an axis point.

Assume that each $\left(\beta_{i-1}^*, \beta_i^*\right)$, $i = 1, \ldots, k$, is neither a corner point nor an axis point. In Section 5.3.3, we developed Algorithm 5.8 for calculating the Lagrange multipliers for this situation. Among those Lagrange multipliers, the λ_i are uniquely determined and $\mu_i = 0$. Hence, all $n_1\left(\beta^*_{i-1}, \beta^*_i; \lambda_i, \mu_i\right)$ and $n_2\left(\beta^*_{i-1}, \beta^*_i; \lambda_i, \mu_i\right)$ are uniquely determined. There is only one feasible solution $\mathbf{q}$ of the problem (TS). Algorithm 5.8 have given the value of $n_1\left(\beta^*_{i-1}, \beta^*_i; \lambda_i, \mu_i\right)$ and $n_2\left(\beta^*_{i-1}, \beta^*_i; \lambda_i, \mu_i\right)$. We need only $k-1$ more addition operations to obtain the value of q_i.

Next assume that there are no axis points but that at least one (β_{i-1}, β_i) is a corner point. By Theorem 5.6, this is the situation when all of the β_i are 1 or -1. Without loss of generality, assume that $\beta_i = 1$, for all $i = 0, \ldots, k$. As discussed in Section 5.3.3, in this situation, every pair $\left(n_1\left(\beta^*_{i-1}, \beta^*_i\right), n_2\left(\beta^*_{i-1}, \beta^*_i\right)\right)$ is defined over the feasible set E_i, $i = 1, \ldots, k$.

Since $n_1\left(\beta^*_{i-1}, \beta^*_i\right) \geq 0$ and $n_2\left(\beta^*_{i-1}, \beta^*_i\right) \geq 0$, we have, from the formulation of problem (TS),

$$q_i = \Delta z_i - n_1\left(\beta^*_{i-1}, \beta^*_i\right) \leq \Delta z_i$$

and

$$q_i = \Delta z_{i-1} + n_2\left(\beta^*_{i-2}, \beta^*_{i-1}\right) \geq \Delta z_{i-1}.$$

Consequently, we have the following lemma.

Lemma 5.12 *If $\beta_i = 1$ for any $i = 0, \ldots, k$, then the sequence $\{q_i\}_{i=1}^{k+1}$ is monotonically nondecreasing.*

If $\Delta z_1 + l_1 \geq 0$, then $q_2 = \Delta z_1 + n_2\left(\beta^*_0, \beta^*_1\right) \geq \Delta z_1 + l_1 \geq 0$. On the other hand, in order to guarantee that $q_2 = \Delta z_1 + n_2\left(\beta^*_0, \beta^*_1\right) \geq 0$, we take the minimum values of both sides, which gives $\Delta z_1 + l_1 \geq 0$. Hence, $q_2 \geq 0$ if and only if $\Delta z_1 + l_1 \geq 0$. If $q_2 \geq 0$, then, by Lemma 5.12, we have

$$q_i \geq q_2 \geq 0\,.$$

for any $i = 2, \ldots, k$. Similarly, since $q_k = \Delta z_{k-1} + n_2\left(\beta^*_{k-2}, \beta^*_{k-1}\right) \leq \Delta z_{k-1} + u_{k-1}$, it is not difficult to prove that $q_k \leq 0$ if and only if $\Delta z_{k-1} + u_{k-1} \leq 0$. These results are summarized in the following lemma.

Lemma 5.13 *Assume that $\beta_i = 1$ for $i = 0, \ldots, k$. All $q_i \geq 0$, $i = 2, \ldots, k$, if and only if $\Delta z_1 + l_1 \geq 0$. All $q_i \leq 0$, $i = 2, \ldots, k$, if and only if $\Delta z_{k-1} + u_{k-1} \leq 0$.*

Assume that all $q_i \geq 0$, $i = 2, \ldots, k$. The objective function of problem (TS) becomes

$$\begin{aligned}
& \sum_{i=1}^{k+1} |q_i| \\
= \quad & |\Delta z_0| + \sum_{i=1}^{k-1} \left(\Delta z_i + n_2 \left(\beta_{i-1}^*, \beta_i^* \right) \right) + \Delta z_{k+1} \\
= \quad & \left[|\Delta z_0| + \sum_{i=1}^{k-1} \Delta z_i + \Delta z_{k+1} \right] + \sum_{i=1}^{k-1} n_2 \left(\beta_{i-1}^*, \beta_i^* \right).
\end{aligned}$$

Therefore, solving the problem (TS) is equivalent to solving the following linear program

$$\text{(TS1)} \quad \begin{cases} \min & \sum\limits_{i=1}^{k-1} n_2 \left(\beta_{i-1}^*, \beta_i^* \right) \\ \text{s.t.} & \left(n_1 \left(\beta_{i-1}^*, \beta_i^* \right), n_2 \left(\beta_{i-1}^*, \beta_i^* \right) \right)^T \in E_i, \quad i = 1, \ldots, k, \\ & n_1 \left(\beta_i^*, \beta_{i+1}^* \right) + n_2 \left(\beta_{i-1}^*, \beta_i^* \right) = -b_i, \quad i = 1, \ldots, k-1. \end{cases}$$

Consider $n_2 \left(\beta_{i-1}^*, \beta_i^* \right)$, $i = 1, \ldots, k-1$, as free variables. Let $R_k(t)$ be the minimal objective function value of problem (TS1) with the additional constraint

$$n_2 \left(\beta_0^*, \beta_1^* \right) \geq t. \tag{5.59}$$

The solution of (TS1) is given in the following theorem.

Theorem 5.14 *Assume that $\beta_i = 1$ for $i = 0, \ldots, k$. If all $q_i \geq 0$, $i = 2, \ldots, k$, then the solution of (TS1) is*

$$n_2 \left(\beta_0^*, \beta_1^* \right) = l_1, \tag{5.60}$$

$$n_2 \left(\beta_{i-1}^*, \beta_i^* \right) = \max \left\{ l_i, \frac{1}{2} \left(-b_{i-1} - n_2 \left(\beta_{i-2}^*, \beta_{i-1}^* \right) \right) \right\}. \tag{5.61}$$

Furthermore,

$$R_k(t) = a_k t + C_k \tag{5.62}$$

is a continuous increasing function, where a_k and C_k are real numbers that depend on the l_i and u_i but not on t, and

$$\frac{1}{2} \leq a_k \leq 1. \tag{5.63}$$

Proof. We prove this theorem by induction.

When $k = 2$, there is only one free variable $n_2(\beta_0^*, \beta_1^*)$. The linear program (TS1) becomes

$$\begin{cases} \min & n_2(\beta_0^*, \beta_1^*) \\ \text{s.t.} & l_1 \le n_2(\beta_0^*, \beta_1^*) \le u_1. \end{cases}$$

Obviously, $R_2(t) = t$, and the optimal solution is $n_2(\beta_0^*, \beta_1^*) = l_1$.

When $k = 3$, there are two free variables $n_2(\beta_0^*, \beta_1^*)$ and $n_2(\beta_1^*, \beta_2^*)$. For any fixed $n_2(\beta_0^*, \beta_1^*) = t$, where $t \in [l_1, u_1]$, the corresponding value of $n_1(\beta_1^*, \beta_2^*)$ is $-b_1 - t$ by (TS1). From the definition of E_i, we know that

$$n_2(\beta_1^*, \beta_2^*) \ge \max\left\{l_2, -\frac{1}{2}b_1 - \frac{1}{2}t\right\}.$$

Let t_0 be the real number such that $l_2 = -\frac{1}{2}b_1 - t_0$, i.e.,

$$t_0 = -b_1 - 2l_2. \tag{5.64}$$

If $t \ge t_0$, then $l_2 \ge -\frac{1}{2}b_1 - \frac{1}{2}t$ and $\min n_2(\beta_1^*, \beta_2^*) = l_2$, which is independent of t. Hence,

$$\begin{aligned} R_3(t) &= \min_{n_2(\beta_0^*,\beta_1^*) \ge t} \{n_2(\beta_0^*, \beta_1^*) + n_2(\beta_1^*, \beta_2^*)\} \\ &= \min_{n_2(\beta_0^*,\beta_1^*) \ge t} n_2(\beta_0^*, \beta_1^*) + \min n_2(\beta_1^*, \beta_2^*) \\ &= t + l_2, \end{aligned}$$

which is a continuous increasing function over $[t_0, u_1]$. The minimum is achieved at $n_2(\beta_1^*, \beta_2^*) = l_2 = \max\left\{l_2, -\frac{1}{2}b_1 - \frac{1}{2}t\right\}$.

If $t \le t_0$, then $l_2 \le -\frac{1}{2}b_1 - \frac{1}{2}t$ and $\min n_2(\beta_1^*, \beta_2^*) = -\frac{1}{2}b_1 - \frac{1}{2}t$. Hence,

$$\begin{aligned} R_3(t) = &\min_{n_2(\beta_0^*,\beta_1^*) \ge t} \{n_2(\beta_0^*, \beta_1^*) + n_2(\beta_1^*, \beta_2^*)\} \\ = &\min\left\{\min_{t \le n_2(\beta_0^*,\beta_1^*) \le t_0} \{n_2(\beta_0^*, \beta_1^*) + n_2(\beta_1^*, \beta_2^*)\}, \min_{n_2(\beta_0^*,\beta_1^*) \ge t_0} \{n_2(\beta_0^*, \beta_1^*) + n_2(\beta_1^*, \beta_2^*)\}\right\} \\ = &\min\left\{\min_{t \le n_2(\beta_0^*,\beta_1^*) \le t_0} \{n_2(\beta_0^*, \beta_1^*) + n_2(\beta_1^*, \beta_2^*)\}, t_0 + l_2\right\}, \end{aligned}$$

where

$$\begin{aligned}
& \min_{t \le n_2\left(\beta_0^*,\beta_1^*\right) \le t_0} \left\{ n_2\left(\beta_0^*,\beta_1^*\right) + n_2\left(\beta_1^*,\beta_2^*\right) \right\} \\
= & \min_{t \le n_2\left(\beta_0^*,\beta_1^*\right) \le t_0} \left\{ n_2\left(\beta_0^*,\beta_1^*\right) \right\} - \frac{1}{2}b_1 - \frac{1}{2}t \\
= & \frac{1}{2}t - \frac{1}{2}b_1.
\end{aligned}$$

Note that $\frac{1}{2}t - \frac{1}{2}b_1$ is a continuous increasing function of t over $[l_1, t_0]$ and that

$$\begin{aligned}
& \left.\frac{1}{2}t - \frac{1}{2}b_1\right|_{t=t_0} \\
& = \frac{1}{2}\left(-b_1 - 2l_2\right) - \frac{1}{2}b_1 \\
& = -b_1 - l_2 \\
& = t_0 + l_2.
\end{aligned}$$

Therefore, when $t \le t_0$,

$$\begin{aligned}
R_3(t) &= \min\left\{ \frac{1}{2}t - \frac{1}{2}b_1, t_0 + l_2 \right\} \\
&= \frac{1}{2}t - \frac{1}{2}b_1.
\end{aligned}$$

The minimum is achieved at $n_2\left(\beta_1^*,\beta_2^*\right) = -\frac{1}{2}b_1 - \frac{1}{2}t = \max\left\{l_2, -\frac{1}{2}b_1 - \frac{1}{2}t\right\}$.

In summary,

$$R_3(t) = \begin{cases} t + l_2, & \text{if } t \ge t_0, \\ \frac{1}{2}t - \frac{1}{2}b_1, & \text{if } t \le t_0. \end{cases}$$

$R_3(t)$ is a continuous increasing function over $[l_1, u_1]$. The coefficient a_3 is either 1 or $\frac{1}{2}$. The minimum is achieved at $n_2\left(\beta_1^*,\beta_2^*\right) = \max\left\{l_2, -\frac{1}{2}b_1 - \frac{1}{2}t\right\}$.

From the definition of $R_3(t)$, it is clear that the minimum objective function value of (TS1) when $k = 3$ is just $R_3(l_1)$, and that the optimal solution is $n_2\left(\beta_0^*,\beta_1^*\right) = l_1$ and $n_2\left(\beta_1^*,\beta_2^*\right) = \max\left\{l_2, -\frac{1}{2}b_1 - \frac{1}{2}t\right\}$. The theorem holds when $k = 3$.

Assume that the conclusions of the theorem are true when $k = j$. Consider the case $k = j + 1$.

For an arbitrary $n_2\left(\beta_0^*,\beta_1^*\right) = t$ such that $t \in [l_1, u_1]$, the lowest value of $n_2\left(\beta_1^*,\beta_2^*\right)$ is given by

$$n_2\left(\beta_1^*,\beta_2^*\right) \ge \max\left\{ l_2, -\frac{1}{2}b_1 - \frac{1}{2}t \right\}.$$

If $t \geq t_0$, then $l_2 \geq -\frac{1}{2}b_1 - \frac{1}{2}t$ and $\min n_2\left(\beta_1^*, \beta_2^*\right) = l_2$. Applying the induction assumption to the last j variables $n_2\left(\beta_1^*, \beta_2^*\right), \ldots, n_2\left(\beta_{k-2}^*, \beta_{k-1}^*\right)$ results in

$$\begin{aligned}
\sum_{i=1}^{k-1} n_2\left(\beta_{i-1}^*, \beta_i^*\right) &\geq t + \min_{n_2\left(\beta_1^*, \beta_2^*\right) \geq l_2} R_j\left(n_2\left(\beta_1^*, \beta_2^*\right)\right) \\
&= t + R_j\left(l_2\right) \\
&= t + a_j l_2 + C_j, \quad t \geq t_0.
\end{aligned}$$

Consequently,

$$R_{j+1}(t) = t + a_j l_2 + C_j, \quad t \geq t_{0.}$$

This is a continuous increasing function of t over $[t_0, u_1]$. The minimum is achieved at $n_2\left(\beta_1^*, \beta_2^*\right) = l_2 = \max\left\{l_2, -\frac{1}{2}b_1 - \frac{1}{2}t\right\}$.

If $t \leq t_0$, then $l_2 \leq -\frac{1}{2}b_1 - \frac{1}{2}t$ and $\min n_2\left(\beta_1^*, \beta_2^*\right) = -\frac{1}{2}b_1 - \frac{1}{2}t$. In analogy to what is above,

$$\begin{aligned}
\sum_{i=1}^{k-1} n_2\left(\beta_{i-1}^*, \beta_i^*\right) &\geq t + \min_{n_2\left(\beta_1^*, \beta_2^*\right) \geq -\frac{1}{2}b_1 - \frac{1}{2}t} R_j\left(n_2\left(\beta_1^*, \beta_2^*\right)\right) \\
&= t + R_j\left(-\frac{1}{2}b_1 - \frac{1}{2}t\right) \\
&= \left(1 - \frac{a_j}{2}\right)t - b_1 C_j, \quad t \leq t_0.
\end{aligned}$$

The minimum is achieved at $n_2\left(\beta_1^*, \beta_2^*\right) = -\frac{1}{2}b_1 - \frac{1}{2}t = \max\left\{l_2, -\frac{1}{2}b_1 - \frac{1}{2}t\right\}$. The induction assumption $\frac{1}{2} \leq a_j \leq 1$ assures that $\frac{1}{2} \leq 1 - \frac{a_j}{2} \leq 1$. Hence, $\left(1 - \frac{a_j}{2}\right)t - b_1 C_j$ is a piecewise increasing function of t over $[l_1, t_0]$. The other induction assumption that $R_j(\cdot)$ is continuous guarantees that $\left(1 - \frac{a_j}{2}\right)t - b_1 C_j$ is a continuous function over $[l_1, t_0]$. Therefore, it must be a continuous increasing function over $[l_1, t_0]$. Furthermore, from (5.64) we know that at point t_0

$$-\frac{1}{2}b_1 - \frac{1}{2}t\bigg|_{t=t_0} = l_2,$$

which implies that $\left(1 - \frac{a_j}{2}\right)t - b_1 C_j$ and $t + a_j l_2 + C_j$ are continuous at t_0. As a consequence,

$$R_{j+1}(t) = \left(1 - \frac{a_j}{2}\right)t - b_1 C_j, \quad t \leq t_{0.}$$

Finally, the optimal objective function value of (TS1) is obviously $R_{j+1}(l_1)$. This minimal value is achieved at $n_2\left(\beta_0^*, \beta_1^*\right) = l_1$. The theorem holds for $k = j + 1$. ■

Assume that $q_i \leq 0$ for all $i = 2, \ldots, k$. The objective function of problem (TS) becomes

$$\begin{aligned}\sum_{i=1}^{k+1} |q_i| &= -\Delta z_0 - \sum_{i=2}^{k} \left(\Delta z_i - n_1\left(\beta_{i-1}^*, \beta_i^*\right)\right) + |\Delta z_{k+1}| \\ &= \left[-\Delta z_0 - \sum_{i=2}^{k} \Delta z_i + |\Delta z_{k+1}|\right] + \sum_{i=2}^{k} n_1\left(\beta_{i-1}^*, \beta_i^*\right).\end{aligned}$$

In the definition of E_i, $i = 2, \ldots, k$, $n_1\left(\beta_{i-1}^*, \beta_i^*\right)$ and $n_2\left(\beta_{i-1}^*, \beta_i^*\right)$ occur symmetrically. If we denote

$$\tilde{l}_i = -b_{i-1} - u_{i-1}, \quad \tilde{u}_i = -b_{i-1} - l_{i-1}, \quad i = 2, \ldots, k,$$

then E_i, $i = 2, \ldots, k-1$, can be restated in the form

$$\begin{aligned} E_i = \Big\{ & \, n_1\left(\beta_{i-1}^*, \beta_i^*\right) \in \left[\tilde{l}_i, \tilde{u}_i\right], \\ & n_2\left(\beta_{i-1}^*, \beta_i^*\right) \in \left[-b_{i-1} - \tilde{u}_{i+1}, -b_{i-1} - \tilde{l}_{i+1}\right], \\ & \tfrac{1}{2} n_1\left(\beta_{i-1}^*, \beta_i^*\right) \leq n_2\left(\beta_{i-1}^*, \beta_i^*\right) \leq 2 n_1\left(\beta_{i-1}^*, \beta_i^*\right) \Big\}. \end{aligned} \tag{5.65}$$

Therefore, solving the problem (TS) is equivalent to solving the following linear program

$$\text{(TS2)} \quad \begin{cases} \min & \sum_{i=2}^{k} n_1\left(\beta_{i-1}^*, \beta_i^*\right) \\ \text{s.t.} & \left(n_1\left(\beta_{i-1}^*, \beta_i^*\right), n_2\left(\beta_{i-1}^*, \beta_i^*\right)\right)^T \in E_i, \quad i = 1, \ldots, k. \\ & n_1\left(\beta_i^*, \beta_{i+1}^*\right) + n_2\left(\beta_{i-1}^*, \beta_i^*\right) = -b_i, \quad i = 2, \ldots, k. \end{cases}$$

where $n_1\left(\beta_{k-1}^*, \beta_k^*\right), \ldots, n_1\left(\beta_1^*, \beta_2^*\right)$ are considered free variables.

Let $\tilde{R}_k(t)$ be the minimal objective function value of problem (TS2) under the additional constraint

$$n_1\left(\beta_{k-1}^*, \beta_k^*\right) \geq t. \tag{5.66}$$

By the logic used in the proof of Theorem 5.14, it is not difficult to prove the following theorem.

Theorem 5.15 *Assume that $\beta_i = 1$ for $i = 0, \ldots, k$. If all $q_i \leq 0$, $i = 2, \ldots, k$, then the solution of (TS2) is*

$$n_1\left(\beta_{k-1}^*, \beta_k^*\right) = \tilde{l}_k = -b_{k-1} - u_{k-1}, \tag{5.67}$$

$$\begin{aligned} n_1\left(\beta_{i-1}^*, \beta_i^*\right) &= \max\left\{\tilde{l}_i, \frac{1}{2}\left(-b_i - n_1\left(\beta_i^*, \beta_{i+1}^*\right)\right)\right\} \\ &= \max\left\{-b_{i-1} - u_{i-1}, \frac{1}{2}\left(-b_i - n_1\left(\beta_i^*, \beta_{i+1}^*\right)\right)\right\}. \end{aligned} \tag{5.68}$$

Furthermore,

$$\tilde{R}_k\left(t\right) = \tilde{a}_k t + \tilde{C}_k \tag{5.69}$$

is a continuous increasing function, where $\tilde{a}_k$ and $\tilde{C}_k$ are real numbers depending on the l_i and u_i but not on t, and

$$\frac{1}{2} \leq \tilde{a}_k \leq 1. \tag{5.70}$$

If the q_i have different signs, then by Lemma 5.12 there must exist an i_0 such that

$$\Delta z_{i_0} + l_{i_0} \leq 0 \quad \text{and} \quad \Delta z_{i_0} + u_{i_0} \geq 0.$$

For all $i > i_0$, $q_{i+1} \geq 0$ and, for all $i < i_0$, $q_{i+1} \leq 0$. Then solving problem (TS) is equivalent to solving the linear program

$$\text{(TS3)} \quad \begin{cases} \min & \sum\limits_{i=2}^{i_0} n_1\left(\beta_{i-1}^*, \beta_i^*\right) + |q_{i_0+1}| + \sum\limits_{i=i_0+1}^{k-1} n_2\left(\beta_{i-1}^*, \beta_i^*\right) \\ \text{s.t.} & \left(n_1\left(\beta_{i-1}^*, \beta_i^*\right), n_2\left(\beta_{i-1}^*, \beta_i^*\right)\right)^T \in E_i, \quad i = 1, \ldots, k. \\ & n_1\left(\beta_i^*, \beta_{i+1}^*\right) + n_2\left(\beta_{i-1}^*, \beta_i^*\right) = -b_i, \quad i = 2, \ldots, k. \end{cases}$$

For the non-negative q_{i+1}, $n_2\left(\beta_{i-1}^*, \beta_i^*\right)$ is considered free variable, $i = i_0+1, \ldots, k-1$. For the non-positive q_i, $n_1\left(\beta_{i-1}^*, \beta_i^*\right)$ is considered free variable, $i = 2, \ldots, i_0$. There is still one degree of freedom. Depending on the sign of q_{i_0+1}, we can consider either $n_1\left(\beta_{i_0}^*, \beta_{i_0+1}^*\right)$ or $n_2\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right)$ to be a free variable as described in what follows.

For any fixed value $n_2\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right) = t$ with

$$-\Delta z_{i_0} \leq t \leq u_{i_0}, \tag{5.71}$$

we know that $q_{i_0+1} = \Delta z_{i_0} + n_2\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right) \geq 0$. Consequently, $|q_{i_0+1}| = \Delta z_{i_0} + n_2\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right)$. Considering $n_2\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right)$ to be a free variable and applying the conclusions of Theorem 5.14 to $q_{i_0+1}, \ldots, q_k$, we find that the minimum value of $\sum_{i=i_0}^{k-1} n_2\left(\beta_{i-1}^*, \beta_i^*\right)$ is

$$R_{k-i_0+1}\left(t\right) = a_{k-i_0+1} t + C_{k-i_0+1}. \tag{5.72}$$

For $n_2\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right) = t$, the value of $n_1\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right)$ is bounded below by

$$n_1\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right) \geq \max\left\{\tilde{l}_{i_0}, \frac{1}{2}t\right\}.$$

Applying the conclusions of Theorem 5.15 to $q_{i_0}, \ldots, q_2$, the minimum value of $\sum_{i=2}^{i_0} n_1\left(\beta_{i-1}^*, \beta_i^*\right)$ is either

$$\tilde{R}_{i_0}\left(\tilde{l}_{i_0}\right) = \tilde{a}_k \tilde{l}_{i_0} + \tilde{C}_k \tag{5.73}$$

or

$$\tilde{R}_{i_0}\left(\frac{1}{2}t\right) = \frac{1}{2}\tilde{a}_k t + \tilde{C}_k. \tag{5.74}$$

All functions defined by (5.72), (5.73), and (5.74) are continuous increasing functions of t. Hence, the optimal value of (TS3) when $q_{i_0+1} \geq 0$ is given by

$$\Delta z_{i_0} + \min_{-\Delta z_{i_0} \leq t \leq u_{i_0}} \left\{ R_{k-i_0+1}(t) + \min\left\{ \tilde{R}_{i_0}\left(\tilde{l}_{i_0}\right), \tilde{R}_{i_0}\left(\frac{1}{2}t\right) \right\} \right\},$$

which is again a continuous increasing function of t. Therefore, the optimal value of t is $-\Delta z_{i_0}$ and hence $|q_{i_0+1}| = 0$.

For any fixed value $n_2\left(\beta_{i_0-1}^*, \beta_{i_0}^*\right) = t$ such that

$$l_i \leq t \leq -\Delta z_{i_0}, \tag{5.75}$$

the value of q_{i_0+1} is negative. We consider $n_1\left(\beta_{i_0}^*, \beta_{i_0+1}^*\right)$ to be a free variable. Then the value of $n_1\left(\beta_{i_0}^*, \beta_{i_0+1}^*\right)$ is $-b_{i_0+1} - t$ and the lower bound of $n_2\left(\beta_{i_0}^*, \beta_{i_0+1}^*\right)$ is $\max\left\{l_{i_0+1}, \frac{1}{2}\left(-b_{i_0+1} - t\right)\right\}$. The conclusions of Theorem 5.14 can be applied to $q_{i_0+1}, \ldots, q_2$ to obtain

$$\tilde{R}_{i_0+1}\left(-b_{i_0+1} - t\right) = \tilde{a}_k\left(-b_{i_0+1} - t\right) + \tilde{C}_k.$$

The conclusions of Theorem 5.15 can be applied to $q_{i_0+2}, \ldots, q_k$ to obtain either

$$R_{k-i_0}\left(l_{i_0+1}\right) = a_{k-i_0} l_{i_0+1} + C_{k-i_0}$$

or

$$R_{k-i_0}\left(\frac{1}{2}\left(-b_{i_0+1} - t\right)\right) = \frac{1}{2} a_{k-i_0}\left(-b_{i_0+1} - t\right) + C_{k-i_0}.$$

These are all continuous decreasing functions of t. As a consequence, when $q_{i_0+1} \leq 0$, the optimal objective function value of (TS3) is achieved at $t = -\Delta z_{i_0}$, i.e., at $|q_{i_0+1}| = 0$. This result is summarized in the following theorem.

Theorem 5.16 *Assume that $\beta_i = 1$ for some i, $i = 0, \ldots, k$. If there exists i_0 such that*

$$\Delta z_{i_0} + l_{i_0} \leq 0, \quad \textit{and} \quad \Delta z_{i_0} + u_{i_0} \geq 0,$$

then the optimal primal solution is achieved at

$$|q_{i_0+1}| = 0. \tag{5.76}$$

When $\beta_i = -1$ for $i = 0, \ldots, k$, similar results can be developed without difficulty. We state the results in this situation without proof.

Lemma 5.17 *Assume that $\beta_i = -1$ for $i = 0, \ldots, k$. Then $q_i \geq 0$ for all i, $i = 2, \ldots, k$, if and only if $\Delta z_{k-1} + l_{k-1} \geq 0$. $q_i \leq 0$ for all i, $i = 2, \ldots, k$, if and only if $\Delta z_1 + u_1 \leq 0$.*

Let $P_k(t)$ be the optimal value of $\max \sum_{i=1}^{k-1} n_2\left(\beta_{i-1}^*, \beta_i^*\right)$ over all feasible $n_2\left(\beta_{i-1}^*, \beta_i^*\right)$ under the additional constraint

$$n_2\left(\beta_0^*, \beta_1^*\right) \leq t. \tag{5.77}$$

Theorem 5.18 *Assume that $\beta_i = -1$ for $i = 0, \ldots, k$. If all $q_i \leq 0$, $i = 2, \ldots, k$, then, at the primal optimal solution, we have*

$$n_2\left(\beta_0^*, \beta_1^*\right) = u_1, \tag{5.78}$$

$$n_2\left(\beta_{i-1}^*, \beta_i^*\right) = \min\left\{u_i, \frac{1}{2}\left(-b_{i-1} - n_2\left(\beta_{i-2}^*, \beta_{i-1}^*\right)\right)\right\}. \tag{5.79}$$

Furthermore,

$$P_k(t) = a_k t + C_k$$

is a continuous increasing function, where a_k and C_k are real numbers that depend on the l_i and u_i but not on t, and

$$\frac{1}{2} \leq a_k \leq 1.$$

Let $\tilde{P}_k(t)$ be the optimal value of $\max \sum_{i=2}^{k} n_1\left(\beta_{i-1}^*, \beta_i^*\right)$ over all feasible $n_1\left(\beta_{i-1}^*, \beta_i^*\right)$ under the additional constraint

$$n_1\left(\beta_{k-2}^*, \beta_{k-1}^*\right) \leq t. \tag{5.80}$$

Theorem 5.19 *Assume that $\beta_i = -1$ for $i = 0, \ldots, k$. If all $q_i \geq 0$, $i = 2, \ldots, k$, then, at the primal optimal solution, we have*

$$n_1\left(\beta_{k-2}^*, \beta_{k-1}^*\right) = \tilde{u}_{k-1}, \tag{5.81}$$

$$n_1\left(\beta_{i-1}^*, \beta_i^*\right) = \min\left\{\tilde{u}_i, \frac{1}{2}\left(-b_i - n_1\left(\beta_i^*, \beta_{i+1}^*\right)\right)\right\}. \tag{5.82}$$

Furthermore,

$$\tilde{P}_k(t) = \tilde{a}_k t + \tilde{C}_k$$

is a continuous increasing function, where $\tilde{a}_k$ and $\tilde{C}_k$ are real numbers that depend on the l_i and u_i but not on t, and

$$\frac{1}{2} \leq \tilde{a}_k \leq 1.$$

Theorem 5.20 *Assume that $\beta_i = -1$ for $i = 0, \ldots, k$. If there exists i_0 such that*

$$\Delta z_{i_0} + l_{i_0} \geq 0 \quad \text{and} \quad \Delta z_{i_0} + u_{i_0} \leq 0,$$

then the optimal primal solution is achieved at

$$|q_{i_0+1}| = 0.$$

5.5 Convergence and complexity

In the algorithm for solving the dual problem, each feasible solution strictly decreases the objective function value. Hence, once a segmentation is checked, it will never appear again. There are a total of 2^{n-2} possible segmentations. Therefore, the algorithm must stop in a finite number of steps. In summary, we have the following theorem.

Theorem 5.21 (Convergence of Algorithm 5.2) *Algorithm 5.2 stops in a finite number of iterations.*

In our empirical experience, the segmentation changes only a few times. For each segment, it takes no more than five iterations to find the local minimizer. Every subalgorithm needs only simple algebraic operations and the number of operations is about the order of the number of knots in the segment. A thorough complexity analysis will be conducted in the future.

All of examples in this thesis are calculated by the algorithm described in this chapter.

5.6 Conclusions

An algorithm based on the geometric programming formulation of univariate cubic L_1 splines is developed in this chapter.

Compared with the widely used discretization approach, this algorithm has several advantages. First, the result obtained by the algorithm developed here is a true optimal solution, while that obtained by the discretization approach is generically only an approximation. Second, this algorithm requires only simple algebraic operations; there is no need to do any matrix operations. Third, this algorithm enables highly parallel processing. At last, it is expected that this algorithm is numerically more stable since it does not require calculation of the inverse of a matrix.

Chapter 6

Conclusion and future research

Cubic L_1 splines have been shown by empirical experience to be promising for shape-preserving multi-scale C^1-smooth interpolating and smoothing. However, theoretical support and efficient algorithms for finding exact solutions have been lacking. This dissertation is directed to correct these deficiencies.

Our approach is to formulate the cubic L_1 splines as a geometric programming problem. The mathematical framework includes a geometric dual program and a linear program for dual-to-primal transformation. The coefficients of a cubic L_1 spline are implicitly determined by a dual optimal solution. We are able to take advantage of this relationship to show some shape-preserving properties of cubic L_1 splines for multi-scale data. It also becomes clear when a regularization function is required to make a cubic L_1 spline unique. In particular, we have shown that cubic L_1 splines preserve linearity over more than three consecutive sub-intervals, and a cubic L_1 spline will never cross the curve of the piecewise linear function for given convex data. Therefore, cubic L_1 splines eliminate oscillations for both linear and convex data.

We have also developed a continuum-based algorithm for obtaining an exact solution for cubic L_1 splines. We take advantage of the special structure of the geometric programming model to decompose the problem into several independent small-sized sub-problems. Then a specialized active set algorithm is developed to solve the sub-problems. The algorithm requires only simple algebraic operations. It is numerically stable and highly parallelizable.

In the future, current research can be extended in the following directions:

(i) The algorithm developed in Chapter 5 appears to take only a few iterations to obtain an optimal solution. We conjecture that its average performance is in order of n, where n is the number of knots. A computational complexity analysis of the number of algebraic operations is needed. An implementation for real applications and a head-to-head comparison with the discretization method would also be extremely interesting.

(ii) This dissertation has proven some shape-preserving properties of univariate cubic L_1 splines and developed a continuum-based algorithm for finding them. We would like to establish a geometric programming model for bi-variate cubic L_1 splines for theoretical analysis and algorithm development.

(iii) We have established a geometric programming model for univariate cubic L_1 smoothing splines in Chapter 3. Parallel to Chapter 5, we would like to develop a continuum-based algorithm and extend the model and algorithm for bi-variate cubic L_1 smoothing splines.

Bibliography

[1] Bojanov, B.D., Hakopian, H.A. and Sahakian, A.A., *Spline Functions and Multivariate Interpolations*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[2] Cheng, H., Fang, S.-C. and Lavery, J.E., "Univariate cubic L_1 splines - a geometric programming approach," *Mathematical Methods of Operations Research*, Vol. 56, issue 2, pp 197-229, 2002.

[3] Clarke, F.H., *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.

[4] Dierckx, P., *Curve and Surface Fitting with Splines*, Oxford, Clarendon, New York, 1993.

[5] Dreyfus, S.E. and Law, A.M., *The Art and Theory of Dynamic Programming*, Academic Press, New York, 1977.

[6] Duffin, R.J., Peterson, E.L. and Zener, C., *Geometric Programming: Theory and Application*, John Wiley & Sons, New York, 1967.

[7] Fang, S.-C. and Puthenpura, S., *Linear Optimization and Extensions: Theory and Algorithms*, Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[8] Farin, G.E., *Curves and Surfaces for Computer Aided Geometric Design : A Practical Guide*, 4th ed., Academic Press, San Diego, 1997.

[9] Fisher, S. and Jerome, J., *Minimum Norm Extremals in Function Spaces with Applications to Classical and Modern Analysis*, Springer-Verlag, Berlin, Heidelberg, 1975.

[10] Fletcher R., *Practical Methods of Optimization, Vol. 1. Unconstrained Optimization*, John Wiley & Sons, Chichester, 1980.

[11] Gill, P.E., Murray, W. and Wright, M.H., *Practical Optimization*, Academic Press, London, 1987.

[12] Hoschek, J. and Lasser, D., *Fundamentals of Computer Aided Geometric Design*, translated from the Germany by Larry L. Schumaker, A K Peters, Wellesley, Massachusetts, 1993.

[13] Kiwiel, K.C., *Methods of Descent for Nondifferentiable Optimization*, Springer-Verlag, Berlin, Heidelberg, 1985.

[14] Lavery, J.E., "Univariate cubic L_p splines and shape-preserving, multiscale interpolation by univariate cubic L_1 splines", *Computed Aided Geometric Design*, Vol. 17, 319-336, 2000.

[15] Lavery, J.E., "Shape-preserving, multiscale fitting of univariate data by cubic L_1 smoothing splines", *Computed Aided Geometric Design*, Vol. 17, 715-727, 2000.

[16] Lavery, J.E., "Shape-preserving, multiscale interpolation by bi- and multivariate cubic L_1 splines", *Computed Aided Geometric Design*, Vol. 18, 321-343, 2001.

[17] Micula, G. and Micula, S., *Handbook of Splines*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[18] Nemhauser, G.L., *Introduction to Dynamic Programming*, John Wiley & Sons, New York, 1966.

[19] Nürnberger, G., *Approximation by Spline Functions*, Springer-Verlag, Berlin, Heidelberg, 1989.

[20] Peña, J.M. (ed.), *Shape Preserving Representations in Computer-Aided Geometric Design*, Nova Science Publishers, Commack, New York, 1999.

[21] Peterson, E.L., "Symmetric duality for generalized unconstrained geometric programming", *SIAM J. Appl. Math.*, Vol. 19, 487-526, 1970.

[22] Peterson, E.L., "Geometric programming", *SIAM Review*, Vol. 18, 1-51, 1976.

[23] Peterson, E.L., "The duality between suboptimization and parameter deletion", *Mathematics of Operations Research*, Vol. 2, 311-319, 1977.

[24] Rockafellar, R.T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[25] Schumaker, L.L., *Spline Functions: Basic Theory*, John Wiley & Sons, New York, 1981.

[26] Scott, S. and Fang, S.-C. (edts), *Geometric Programming*, Annals of Operations Research, Vol. 105, 2001.

[27] Shor, N.Z., *Minimization Methods for Non-Differentiable Functions*, translated from the Russian by K. C. Kiwiel and A. Ruszczyński, Springer-Verlag, Berlin, Heidelberg, 1985.

[28] Späth, H., *One Dimensional Spline Interpolation Algorithms*, A K Peters, Wellesley, Massachusetts, 1995.

[29] Späth, H., *Two Dimensional Spline Interpolation Algorithms*, A K Peters, Wellesley, Massachusetts, 1995.

[30] Zoutendijk, G., *Methods of Feasible Directions: A Study in Linear and Nonlinear Programming*, Elsevier, Amsterdam, 1960.