Tridiagonally Separable Programs: Unconstrained Duals and Applications in Convex Spline Approximation*

Jochen W. Schmidt
Institute of Numerical Mathematics,
Technical University of Dresden, D-01062 Dresden,
Germany
e-mail: jschmidt@math.tu-dresden.de

Abstract

If the Fenchel conjugates are everywhere finite, convex tridiagonally separable programs can be dualized such that the concave duals become unconstrained. By means of the return-formula the solution of the primal is explicitly computed if a solution of the dual is known. Therefore, from a numerical point of view, dualization is very attractive in these optimization problems. The dual approach is applied to some types of convex spline approximation where, for selecting a preferable spline solution, the curvature or other more simple functionals are minimized.

Classification: 90C25, 65D07, 41A15

Key Words: Unconstrained non-smooth duals, Fenchel conjugates, quadratic splines on refined grids, curvature minimization, fit-and-modify strategy.

1 Introduction

Introduction
In univariate shape preserving spline interpolation, we are frequently led to the following problem. For given nonempty closed convex sets $W_1, W_2, \ldots, W_n \in \mathbb{R}^{2\ell}$, assure the existence of vectors $p_0, p_1, \ldots, p_n \in \mathbb{R}^\ell$ satisfying

$$(p_{i-1}, p_i) \in W_i, \quad i = 1, 2, \ldots, n. \tag{1.1}$$

*Extended version of a lecture given at the 3rd International Conference on Approximation and Optimization in the Caribbean, Puebla, October 1995.

The staircase algorithm [10], [2] provides a existence test being necessary and sufficient. If solvable at all, the number of solutions of (1.1) is infinite in general, and so is the number of spline approximants. In order to find a preferable spline solution it has become more and more usual to introduce a choice functional, e.g. the Holladay functional [6]. But there are also other functionals in use. In this way an optimization problem arises which often has the very special structure

$$P: \qquad \text{minimize} \left\{ \sum_{i=1}^{n} F_i\,(p_{i-1}, p_i) \; : \; (p_{i-1}, p_i) \in W_i, , \quad i = 1, \ldots, n \right\}, \tag{1.2}$$

where $F_1, \ldots, F_n \; : \; I\!R^{2\ell} \to I\!R^1$ are convex functions. (Like in (1.1) and (1.2), for $p, q \in I\!R^\ell$ we will always write (p, q) instead of the more correct notation $(p^T, q^T)^T$.) The program P is called a tridiagonally separable one. The reason is the block-tridiagonality of the Hessian of the objective function, and the adapted form of the constraints.

For solving program P effectively one should utilize its structure. Direct methods with this property seem not to be developed. This may be feeled as a lack. There is an interest in methods which are both fast and running in the feasible domain. Because of the background of the problems the feasibility is more important than a very accurate optimal value with respect to the just introduced choice function.

In the present paper we offer dual methods, though the desired feasibility does not appear. On the other hand, in our problems we obtain dual programs D which have the advantage to be unconstrained. Thus the effective methods of unconstrained optimization apply. In addition, by means of the return-formula the solution of the original program P can be explicitly computed if a solution of the dual D is known. For performing this dual approach we mainly need the Fenchel conjugates which are determined by programs of the lower dimension 2ℓ.

To our knowledge, a first example of a dual algorithm for a partially separable program was derived in [1] where convex interpolation with cubic splines is treated. Further we refer to [11] and [4]. A theoretical foundation of the dual methods under consideration can be found in the papers [12] and [8]. In the latter paper also several examples of applications are presented. In [3] the duality properties of P and its dual are shown to be consequences of corresponding statements on more simple decomposable programs [5], [7].

Now we are concerned with recent results in univariate shape preserving spline approximation. It has turned out that certain splines on refined grids allow some of the approximation problems to be always solvable if the additional nodes are suitably placed. A first step in this direction was done in the paper [14] Again we are led to tridiagonally separable programs P, and there is a real interest to prepare also the now arising ones for applying dual methods. In particular, the Fenchel conjugates are required.

2 Duality Statements for Tridiagonally Separable Programs

Duality Statements for Tridiagonally Separable Programs
We consider the following tridiagonally separable program P that is somewhat more general then (1.2),

$$\begin{aligned} &\text{minimize } \sum_{i=1}^{n} F_i(p_{i-1}, p_i) \\ &\text{s.t. } \sum_{i=1}^{n} (B_i\, p_{i-1} + C_i\, p_i) = b, \quad (p_{i-1}, p_i) \in W_i,\ i = 1, \dots, n. \end{aligned} \tag{2.1}$$

As before the functions $F_i : I\!R^{2\ell} \to I\!R^1$ are assumed to be convex and the sets $W_i \in I\!R^{2\ell}$ to be nonempty, closed, and convex, but not necessarily compact. Further let B_i and C_i be (m, ℓ)-matrices, and $b \in I\!R^m$. Program (1.2) is obtained by setting $B_i = C_i = 0$, $i = 1, \dots, n$ and $b = 0$.

From [8] we recall that a dual program D to (2.1) reads

$$\begin{aligned} &\text{maximize} \quad -\sum_{i=1}^{n} H_i^*(p_{i-1}^* + B_i^T p_{n+1}^*, -p_i^* + C_i^T p_{n+1}^*) + b^T p_{n+1}^* \\ &\qquad\qquad \text{with } p_0^* = p_n^* = 0\,; \end{aligned} \tag{2.2}$$

the Fenchel conjugates $H_1^*, \dots, H_n^*$ needed here are computed by solving the parametric 2ℓ-dimensional programs

$$H_i^*(\xi, \eta) = \sup\, \{\, x^T \xi + y^T \eta - F_i(x, y) \ :\ (x, y) \in W_i \,\}\,,\quad i = 1, \dots, n. \tag{2.3}$$

Of course, if the conjugates are finite everywhere,

$$H_i^*(\xi, \eta) < +\infty \quad \text{for} \quad (\xi, \eta) \in I\!R^{2\ell},\ \ i = 1, \dots, n\,, \tag{2.4}$$

the dual program (2.2) becomes unconstrained.

Next, let the subgradients of the conjugates be nonempty everywhere,

$$\partial H_i^*(\xi, \eta) \neq \emptyset \quad \text{for} \quad (\xi, \eta) \in I\!R^{2\ell},\ \ i = 1, \dots, n. \tag{2.5}$$

This property is equivalent to the existence of maximizers to the programs (2.3) for all $(\xi, \eta) \in \mathbb{R}^{2\ell}$.

Theorem 1 *Let (2.5) be valid. Then program P is solvable if and only if the optimal value* $\sup D$ *of the dual program D is finite, and this in turn holds if and only if the feasible domain of P is nonempty. If in addition to (2.5), the sets $W_1, \ldots, W_n$ are polyhedral then the feasibility of P implies the solvability of programs P and D. Further, all solutions $(p_0, \ldots, p_n)$ of program P are then given by the return-formula*

$$(p_{i-1}, p_i) \in \partial H_i^*(p_{i-1}^* + B_i^T p_{n+1}^*, -p_i^* + C_i^T p_{n+1}^*), \quad i = 1, \ldots, n, \tag{2.6}$$

where $(p_0^*, p_1^*, \ldots, p_{n+1}^*)$ denotes a solution of the dual program D.

We remark that in the present applications the maximizers in the programs (2.3) are always uniquely determined. This implies the conjugates $H_1^*, \ldots, H_n^*$ to be continuously differentiable on $\mathbb{R}^{2\ell}$. Further, the return-formula (2.6) then reads

$$p_{i-1} = \partial_1 H_i^*(p_{i-1}^* + B_i^T p_{n+1}^*, -p_i^* + C_i^T p_{n+1}^*), \tag{2.7}$$

$$p_i = \partial_2 H_i^*(p_{i-1}^* + B_i^T p_{n+1}^*, -p_i^* + C_i^T p_{n+1}^*), \quad i = 1, \ldots, n,$$

and consequently program P becomes uniquely solvable.

Formula (2.7) yields two values for $p_1, \ldots, p_{n-1}$. This gives a very useful possibility to check the accuracy of the dual solution which normally is determined numerically.

In Theorem 1 the sets $W_1, \ldots, W_n$ are assumed to be polyhedral, i.e., they are described by finite sets of affine linear inequalities. In the other case, the corresponding parts of Theorem 1 remain valid if a constraint qualification is added, e.g., the existence of a vector $(p_0, \ldots, p_n)$ with

$$\sum_{i=1}^{n} (B_i p_{i-1} + C_i p_i) = b, \quad (p_{i-1}, p_i) \in \operatorname{ri} W_i, \quad i = 1, \ldots, n. \tag{2.8}$$

For detailed proofs of these assertions we refer to [8], or to the papers mentioned above.

3 Quadratic C^1 Splines in convex Interpolation

Quadratic C^1 Splines in Convex Interpolation For a given data set D_n:

$$(x_i, y_i) \in \mathbb{R}^2, \quad i = 0, \ldots, n, \tag{3.1}$$

defined on the grid $\Delta_n : x_0 < x_1 < \ldots < x_n$, the problem of convex interpolation is to find a convex C^k function s with

$$s(x_i) = y_i, \quad i = 0, \ldots, n. \tag{3.2}$$

In practical models the integer k may be equal 1,2, or even 3. Of course, this problem is solvable only under fair assumptions on D_n. In the univariate case (3.1), relative to the fairness, D_n is readily assumed to be in strictly convex position, namely

$$\tau_1 < \tau_2 < \ldots < \tau_n \tag{3.3}$$

with the slopes $\tau_i = (y_i - y_{i-1})/h_i$ and the step sizes $h_i = x_i - x_{i-1}$.

For convexity preservation quadratic splines s are considered on grids $\tilde{\Delta}_n$ which originate from Δ_n by adding one node

$$\xi_i = \beta_i x_{i-1} + \alpha_i x_i, \quad \alpha_i > 0,\ \beta_i > 0,\ \alpha_i + \beta_i = 1 \tag{3.4}$$

in each subinterval $[x_{i-1}, x_i]$, $i = 1, \ldots, n$. Ratios α_i, β_i with the property (3.4) will be called feasible. With the local variables $u = (x - x_{i-1})/(\alpha_i h_i)$ and $v = (x_i - x)/(\beta_i h_i)$ we now define the splines by

$$s(x) = \eta_i u + y_{i-1}(1-u) + (y_{i-1} - \eta_i + \alpha_i h_i p_{i-1})\, u\,(1-u) \quad \text{on} \quad [x_{i-1}, \xi_i]\,, \tag{3.5}$$

$$s(x) = \eta_i v + y_i(1 - v) + (y_i - \eta_i - \beta_i h_i p_i)\, v\,(1 - v) \quad \text{on} \quad [\xi_i, x_i]\,.$$

Obviously these splines interpolate D_n, and we have $s \in C^1[x_0, x_n]$ if we set

$$\eta_i = \beta_i(y_{i-1} + \alpha_i h_i p_{i-1}/2) + \alpha_1(y_i - \beta_i h_i p_i/2), \quad i = 1, \ldots, n. \tag{3.6}$$

The unknown parameters $p_0, \ldots, p_n$ turn out to be the first order derivatives in the original nodes,

$$p_i = s'(x_i), \quad i = 0, \ldots, n. \tag{3.7}$$

Further, these splines are immediately seen to be convex on $[x_0, x_n]$ if and only if

$$(1 + \alpha_i)p_{i-1} + (1 - \alpha_i)p_i \le 2\tau_i \le \alpha_i p_{i-1} + (2 - \alpha_i)p_i, \quad i = 1, \ldots, n\,. \tag{3.8}$$

Summarizing these considerations we obtain

Proposition 2 *The problem of convex interpolation is solvable with quadratic C^1 splines on refined grids $\tilde{\Delta}_n$ if and only if the system (3.8) of linear inequalities possesses solutions $(p_0, p_1, \ldots, p_n)$.*

Now, as it is shown in [9], system (3.8) is solvable if the ratios are chosen according to

$$\alpha_1 \in (0,1),\ \alpha_n \in (0,1),$$
$$\alpha_i \in (\max\{0, 1-2(\tau_i - \tau_{i-1})/(\tau_{i+1}-\tau_{i-1})\},\, 1),\ i = 2, \ldots, n-1, \quad (3.9)$$
$$\beta_i = 1 - \alpha_i,\ i = 1, \ldots, n;$$

here, of course, the strict convexity (3.3) has to be assumed. Every solution $(p_0, p_1, \ldots, p_n)$ leads via (3.5), (3.6) to an interpolating spline which is convex on $[x_0, x_n]$.

Let us fix the ratios such that (3.9) is satisfied, e.g., by choosing the middles of the intervals. Now, if the preferred interpolant is selected by means of the Holladay functional we are led to the following program

$$\text{minimize} \int_{x_0}^{x_n} s''(x)^2 dx = \sum_{i=1}^{n} \frac{1}{h_i} \left\{ (p_{i-1} - p_i)^2 + \frac{(2\tau_i - p_{i-1} - p_i)^2}{\alpha_i \beta_i} \right\} \quad (3.10)$$

subject to the convexity constraints (3.8).

Of course, this is a tridiagonally separable program (1.2) with $\ell = 1$, the functions

$$F_i(x,y) = \frac{(x-y)^2}{h_i} + \frac{(2\tau_i - x - y)^2}{\alpha_i \beta_i h_i}\,, \quad (3.11)$$

and the sets

$$W_i = \{(x,y) \in I\!R^2 \ :\ (1+\alpha_i)\,x + (1-\alpha_i)\,y \le 2\tau_i \le \alpha_i x + (2-\alpha_i)\,y\}\,. (3.12)$$

Since the Hessians of the quadratic functions (3.11) are positive definite, we have the

Proposition 3 *Feasible programs (3.10) are uniquely solvable, and feasibility is assured by choosing the ratios according to (3.9).*

For solving (3.10) numerically we propose dualization in the sense of program (2.2), and we will provide the necessary Fenchel conjugates in the chapter after the next.

Another way for selecting a preferred convex spline interpolant is the fit-and-modify approach. Here, an initial approximation $(q_0, q_1, \ldots, q_n)$ to the first order derivatives is modified by norm-minimizing in order to fullfil the convexity constraints. The initial estimates can be computed from the data set D_n, e.g., by using the well-known formulae

$$q_i = \frac{h_i \tau_{i+1} + h_{i+1}\tau_i}{h_i + h_{i+1}} \quad i = 1, \ldots, n-1\,, \tag{3.13}$$

$$q_0 = 2\tau_i - q_1\,, \quad q_n = 2\tau_n - q_n\,.$$

Thus in this approach it is suggested to consider the program

$$\begin{array}{l} \text{minimize} \;\; \sum\limits_{i=1}^{n} \{\, \gamma_{i-1}(p_{i-1} - q_{i-1})^2 + \gamma_i (p_i - q_i)^2 \,\} \\ \text{subject to the convexity constraints (3.8).} \end{array} \tag{3.14}$$

The positive weights may be, e.g., $\gamma_0 = \gamma_n = 2$, $\gamma_1 = \ldots = \gamma_{n-1} = 1$. Again we get a tridiagonally separable program (1.2) with $\ell = 1$. Now we have

$$F_i(x, y) = \gamma_{i-1}(x - q_{i-1})^2 + \gamma_i\,(y - q_i)^2, \tag{3.15}$$

while the W_i are defined as in (3.12). Likewise program (3.14) is uniquely solvable if it is feasible. The reason is the positive definiteness of the Hessians of the functions (3.15). From the numerical point of view, the dual program (2.2) to (3.14) is of interest. The Fenchel conjugates necessary for this purpose will be derived in the chapter after the next.

4 Quadratic C^1 Splines in Convex Smoothing

Quadratic C^1 Splines in Convex Smoothing In this chapter, we will turn our attention from interpolation to smooth approximation, again subject to the constraints of convexity. For convenience, the given data set D_n is now denoted by

$$(x_i,\, u_i) \in \mathbb{R}^2, \;\; i = 0, \ldots, n \tag{4.1}$$

defined on Δ_n. In smoothing, the interpolation condition (3.2) is usually replaced by minimizing the Schoenberg functional [13]

$$\kappa \int_{x_0}^{x_n} s''(x)^2 dx + \sum_{i=1}^{n} \delta_{i-1}(s(x_{i-1}) - u_{i-1})^2 + \delta_i (s(x_i) - u_i)^2 \,\}\,. \tag{4.2}$$

The global parameter $\kappa > 0$ and the local ones $\delta_0 > 0, \ldots, \delta_n > 0$ may be given. The functional (4.2) establishes a compromise between interpolation and smooth approximation. Mainly the choice of the global parameter depends on which of these two conflicting goals is considered to be of greater importance.

When applying quadratic splines (3.5), (3.6), in convex smoothing we obtain a tridiagonally separable program (1.2) with $\ell = 2$ which may be written as

$$\begin{aligned} &\text{minimize } \sum_{i=1}^{n} F_i(y_{i-1}, p_{i-1}, y_i, p_i) \\ &s.t. \ (y_{i-1}, p_{i-1}, y_i, p_i) \in W_i\,, \quad i = 1, \ldots, n. \end{aligned} \tag{4.3}$$

The functions F_i and the sets W_i therein are

$$\begin{aligned} F_i(f, x, g, y) = \ & \tfrac{\kappa}{h_i} \left\{ (x-y)^2 + \tfrac{(2(g-f)/h_i - x - y)^2}{\alpha_i \beta_i} \right\} + \\ & + \ \delta_{i-1}(f - u_{i-1})^2 + \delta_i (g - u_i)^2, \end{aligned} \tag{4.4}$$

$$\begin{aligned} W_i = \ & \{(f, x, g, y) \in I\!R^4 : (1+\alpha_i)x + (1-\alpha_i)y \\ & \leq 2(g-f)/h_i \leq \alpha_i x + (2-\alpha_i)\, y\}. \end{aligned} \tag{4.5}$$

The Hessians of these functions F_i are readily seen to be positive definite. Further, the feasible domain is nonempty for all permitted ratios since, e.g., the vector with $y_i = p_i = 0$, $i = 0, \ldots, n$ is feasible. Thus, we are led to the

Proposition 4 *For all feasible ratios $\alpha_1, \beta_1, \ldots, \alpha_n, \beta_n$ the quadratic program (4.3), (4.4), (4.5) arising in convex smoothing is uniquely solvable.*

In view of the fit-and-modify functional (3.14) also the minimization of the following extended functional is of interest in convex smoothing,

$$\begin{aligned} \textstyle\sum_{i=1}^{n} \{\kappa_{i-1}(s'(x_{i-1}) - q_{i-1})^2 + \\ \kappa_i (s'(x_i) - q_i)^2 + \delta_{i-1}(s(x_{i-1}) - u_{i-1})^2 + \delta_i (s(x_i) - u_i)^2\}, \end{aligned} \tag{4.6}$$

with constants $\kappa_i > 0$, $i = 0, \ldots, n$. Again, the values $q_0, \ldots, q_n$ have to be initial estimates of the first order derivatives in the original nodes. In this way we are also led to a program (4.3), now with the functions

$$F_i(f, x, g, y) = \kappa_{i-1}(x - q_{i-1})^2 + \kappa_i (y - q_i)^2 + \delta_{i-1}(f - u_{i-1})^2 + \delta_i (g - u_i)^2, \tag{4.7}$$

while the sets W_i are given in (4.5). It is easily seen the Proposition 4 holds true likewise for the program (4.3), (4.7), (4.5).

The dualization (2.2) of both programs (4.3) for convex smoothing will be treated in the next chapter.

5 Quadratic C^1 Splines in Convex Smoothing

Dualization of the Programs in Convex Approximation
The tridiagonally separable programs (1.2) and (4.3) which arise in the present interpolation and smoothing problems can be effectively solved by means of dualization. The corresponding dual programs (2.2) have the advantage to be unconstrained. In all cases, the dual objective functions are piecewise quadratic. They are non-smooth in the sense that they are from C^1 but not from C^2. The second order derivatives are not existing on sets of zero measure only.

For establishing the duals as well as the return-formulae the Fenchel conjugates (2.3) are needed.

Proposition 5 *The conjugates (2.3) to the functions (3.11) and the sets (3.12) obtained in selecting convex interpolants by means of the Holladay functional are*

$$H_1^*(\xi,\eta) = \tau_i(\xi+\eta) + h_i\, R_i^*(\xi,\eta) \tag{5.1}$$

with $R_i^*(\xi,\eta) =$

$$\begin{cases} (\xi-\eta)^2 + \alpha_i\beta_i(\xi+\eta)^2 & for \quad (1+\beta_i)\xi \le \alpha_i\eta, \beta_i\xi \le (1+\alpha_i)\eta, \\ \beta_i(\beta_i\xi - (1+\alpha_i)\eta)^2 & for \quad (1+\beta_i)\xi \ge \alpha_i\eta, \beta_i\xi \le (1+\alpha_i)\eta, \\ \alpha_i(\alpha_i\eta - (1+\beta_i)\xi)^2 & for \quad (1+\beta_i)\xi \le \alpha_i\eta, \beta_i\xi \ge (1+\alpha_i)\eta, \\ 0 & for \quad (1+\beta_i)\xi \ge \alpha_i\eta, \beta_i\xi \ge (1+\alpha_i)\eta. \end{cases} \tag{5.2}$$

Proof The Lagrangian Φ to program (2.3) with the function (3.11) and the set (3.12) is given by

$$\begin{aligned} \Phi &= -\xi x - \eta y + \tfrac{(x-y)^2}{h_i} + \tfrac{(\tau_i - x - y)^2}{\alpha_i\beta_i h_i} + \\ &\quad +\lambda((1+\alpha_i)x + \beta_i y - 2\tau_i) + \mu(2\tau_i - \alpha_i x - (1+\beta_i)y)\,. \end{aligned} \tag{5.3}$$

The Kuhn-Tucker conditions, being now necessary as well as sufficient, read

$$\phi_x = 0,\ \phi_y = 0,\ \phi_\lambda \le 0,\ \lambda \ge 0,\ \lambda\phi_\lambda = 0,\ \phi_\mu \le 0,\ \mu \ge 0,\ \mu\phi_\mu = 0. \tag{5.4}$$

The linear system of equations with respect to x and y which results from $\phi_x = 0$ and $\phi_y = 0$ is straightforwardly seen to be solved by

$$\begin{aligned} x &= \tau_i + \frac{h_i}{8}(\xi-\eta) + \frac{h_i\alpha_i\beta_i}{8}(\xi+\eta) - \frac{h_i\alpha_i(1+\beta_i)}{4}\lambda - \frac{h_i\beta_i^2}{4}\,\mu, \\ y &= \tau_i - \tfrac{h_i}{8}(\xi-\eta) + \tfrac{h_i\alpha_i\beta_i}{8}(\xi+\eta) + \tfrac{h_i\alpha_i^2}{4}\lambda + \tfrac{h_i\beta_i(1+\alpha_i)}{4}\,\mu. \end{aligned} \tag{5.5}$$

In view of the complementarity conditions in (5.4), four cases have to be considered.

Case 1: Let $\lambda = 0$, $\mu = 0$. The maximizer given by (5.5) with $\lambda = \mu = 0$ is unique. If the property

$$x = \partial_1 H_i^*(\xi, \eta), \quad y = \partial_2 H_i^*(\xi, \eta) \tag{5.6}$$

is taken into account, the first part of formula (5.1), (5.2) is easily verified. The given domain follows from $\phi_\lambda \le 0$ and $\phi_\mu \le 0$.

Case 2: Let $\phi_\lambda = 0$, $\mu = 0$. Now we obtain

$$\lambda = \frac{1}{2}(\xi - \eta) + \frac{\beta_i}{2}(\xi + \eta), \tag{5.7}$$

and therefore (5.5) yields

$$x = \tau_i + \tfrac{h_i \beta_i^3}{8}\,\xi - \tfrac{h_i \beta_i^2 (1+\alpha_i)}{8}\,\eta\,,$$

$$y = \tau_i - \tfrac{h_i \beta_i^2 (1+\alpha_i)}{8}\,\xi + \tfrac{h_i \beta_i (1+\alpha_i)^2}{8}\,\eta\,.$$

By means of (5.6), the second part of the desired result (5.1), (5.2) is immediately seen to be valid. The corresponding domain follows from $\lambda \ge 0$ and $\phi_\mu \le 0$.

Case 3: Let $\lambda = 0$, $\phi_\mu = 0$. This case is treated analogously to Case 2. We get

$$\mu = \frac{1}{2}(\xi - \eta) - \frac{\alpha_i}{2}(\xi + \eta)\,, \tag{5.8}$$

and

$$x = \tau_i + \tfrac{h_i \alpha_i (1+\beta_i)^2}{8}\,\xi - \tfrac{h_i \alpha_i^2 (1+\beta_i)}{8}\,\eta\,,$$

$$y = \tau_i - \tfrac{h_i \alpha_i^2 (1+\beta_i)}{8}\,\xi + \tfrac{h_i \alpha_i^3}{8}\,\eta\,,$$

implying the third part of (5.1), (5.2).

Case 4: Let $\phi_\lambda = 0$, $\phi_\mu = 0$. The resulting linear system is uniquely solved by

$$x = \tau_i\,, \quad y = \tau_i\,,$$

while the Lagrange factors λ and μ are again given by (5.7) and (5.8), respectively. Thus in view of (5.6), the fourth part of formula (5.1), (5.2) follows, and the proof of Proposition 5 is complete. ■

Proposition 6 *The conjugates (2.3) to the functions (3.15) and the sets (3.12) obtained in selecting convex interpolants by means of the fit-and-modify strategy are* $H_i^*(\xi,\eta) =$

$$= \begin{cases} Z_i(\xi,\eta) \;\; for X_i(\xi,\eta) \geq 0, Y_i(\xi,\eta) \leq 0, \\ Z_i(\xi,\eta) - \frac{\gamma_{i-1}\gamma_i}{4\Gamma_{1i}} X_i(\xi,\eta)^2 \;\; for 0 \geq \Gamma_i X_i(\xi,\eta) \geq \Gamma_{1i} Y_i(\xi,\eta), \\ Z_i(\xi,\eta) - \frac{\gamma_{i-1}\gamma_i}{4\Gamma_{2i}} Y_i(\xi,\eta)^2 \;\; for 0 \leq \Gamma_i Y_i(\xi,\eta) \leq \Gamma_{2i} X_i(\xi,\eta), \\ \tau_i(\xi+\eta) \;\; for \Gamma_i X_i(\xi,\eta) \leq \Gamma_{1i} Y_i(\xi,\eta), \Gamma_i Y_i(\xi,\eta) \geq \Gamma_{2i} X_i(\xi,\eta) \end{cases} \tag{5.9}$$

with the abbreviations

$$L_i \;= 2\tau_i - (1+\alpha_i)\, q_{i-1} - \beta_i q_i,\; U_i = 2\tau_i - \alpha_i q_{i-1} - (1+\beta_i)\, q_i,$$

$$\Gamma_i \;= \alpha_i(1+\alpha_i)\,\gamma_i + \beta_i(1+\beta_i)\,\gamma_{i-1},$$

$$\Gamma_{1i} = (1+\alpha_i)^2\gamma_i + \beta_i^2\gamma_{i-1},$$

$$\Gamma_{2i} = \alpha_i^2\gamma_i + (1+\beta_i)^2\gamma_{i-1}, \tag{5.10}$$

$$X_i(\xi,\eta) = 2L_i - \tfrac{(1+\alpha_i)}{\gamma_{i-1}}\,\xi - \tfrac{\beta_i}{\gamma_i}\,\eta,$$

$$Y_i(\xi,\eta) = 2U_i - \tfrac{\alpha_i}{\gamma_{i-1}}\,\xi - \tfrac{1+\beta_i}{\gamma_i}\,\eta,$$

$$Z_i(\xi,\eta) = q_{i-1}\xi + q_i\eta + \tfrac{\xi^2}{4\gamma_{i-1}} + \tfrac{\eta^2}{4\gamma_i}\,.$$

The previous proof carries over immediately. The Lagrangian is now

$$\begin{aligned} \Phi = \;\; & {}^{-}\xi x - \eta y + \gamma_{i-1}(x-q_{i-1})^2 + \gamma_i(y-q_i)^2 + \\ & +\lambda((1+\alpha_i)x + \beta_i y - 2\tau_i) + \mu(2\tau_i - \alpha_i x - (1+\beta_i)y), \end{aligned} \tag{5.11}$$

and the system of linear equations $\phi_x = 0$ and $\phi_y = 0$ has the solution

$$x = q_{i-1} + (\xi - (1+\alpha_i)\,\lambda + \alpha_i\mu)/(2\gamma_{i-1}), \tag{5.12}$$

$$y = q_i + (\eta - \beta_i\lambda + (1+\beta_i)\mu)/(2\gamma_i).$$

Considering the four cases generated by the complementarity conditions in (5.4), we obtain straightforwardly the representation (5.9) of the conjugates.

Proposition 7 *The conjugates (2.3) to the functions (4.4) and the sets (4.5) which arise in convex approximation by means of the Schoenberg functional read*

$$H_i^*(\varrho,\xi,\sigma,\eta) = u_{i-1}\varrho + u_i\sigma + \frac{u_i - u_{i-1}}{h_i}(\xi+\eta) + \frac{h_i}{16\kappa}R_i^*(\xi,\eta) + \tag{5.13}$$

$$+\tfrac{h_i^2\varrho^2-2h_i\varrho(\xi+\eta)+(\xi+\eta)^2}{4h_i^2\delta_{i-1}}+\tfrac{h_i^2\sigma^2+2h_i\sigma(\xi+\eta)+(\xi+\eta)^2}{4h_i^2\delta_i}$$

where the R_i^* are defined in (5.2), $i=1,\ldots,n$.

Proof The Lagrangian to the present program (2.3) is

$$\Phi=-\varrho f-\xi x-\sigma g-\eta y+\tfrac{\kappa}{h_i}(x-y)^2+\tfrac{\kappa}{\alpha_i\beta_i h_i}\left(2\,\tfrac{g-f}{h_i}-x-y\right)^2+$$
$$+\delta_{i-1}(f-u_{i-1})^2+\delta_i(g-u_i)^2+ \tag{5.14}$$
$$+\lambda\left((1+\alpha_i)x+\beta_i y-2\,\tfrac{g-f}{h_i}\right)+\mu\left(2\,\tfrac{g-f}{h_i}-\alpha_i x-(1+\beta_i)y\right).$$

Straightforwardly we obtain from $\Phi_x=0$, $\Phi_y=0$ that

$$x=\frac{g-f}{h_i}+\frac{h_i}{8\kappa}(\xi-\eta)+\frac{\alpha_i\beta_i h_i}{8\kappa}(\xi+\eta)-\frac{\alpha_i(1+\beta_i)h_i}{4\kappa}\,\lambda-\frac{\beta_i^2 h_i}{4\kappa}\,\mu, \tag{5.15}$$
$$y=\tfrac{g-f}{h_i}-\tfrac{h_i}{8\kappa}(\xi-\eta)+\tfrac{\alpha_i\beta_i h_i}{8\kappa}(\xi+\eta)+\tfrac{\alpha_i^2 h_i}{4\kappa}\,\lambda-\tfrac{\beta_i(1+\alpha_i)h_i}{4\kappa}\,\mu,$$

while $\Phi_f=0$, $\Phi_g=0$ yield

$$f=u_{i-1}+\frac{\varrho}{2\delta_{i-1}}-\frac{\xi+\eta}{2\delta_{i-1}h_i}, \tag{5.16}$$
$$g=u_i+\tfrac{\sigma}{2\delta_i}+\tfrac{\xi+\eta}{2\delta_i h_i}.$$

Notice that f and g do not depend on the Lagrange parameters λ and μ. The complementarity conditions

$$\Phi_\lambda\le 0,\ \lambda\ge 0,\ \lambda\Phi_\lambda=0,\ \Phi_\mu\le 0,\ \mu\ge 0,\ \mu\Phi_\mu=0 \tag{5.17}$$

again require to consider four cases. The desired result (5.13) follows after some computations from (5.14) and (5.15) by applying the formulae

$$f=\partial_1 H_i^*(\varrho,\xi,\sigma,\eta),\quad x=\partial_2 H_i^*(\varrho,\xi,\sigma,\eta), \tag{5.18}$$
$$g=\partial_3 H_i^*(\varrho,\xi,\sigma,\eta),\quad y=\partial_4 H_i^*(\varrho,\xi,\sigma,\eta),$$

which are valid for the maximizer of program (2.3). ■

Next, convex smoothing by means of the fit-and-modify functional (4.6) is considered. We have seen that this problem leads to a tridiagonally separable program (1.2) with $\ell=2$, namely to the optimization problem (4.3), (4.7), (4.5). In dualization the conjugates H_i^* to (4.7), (4.5) are of interest. For the sake of brevity we derive the first order derivatives of H_i^*. There are several methods in unconstrained optimization which require derivatives only and no function values.

We use here the abbreviations

$$\begin{aligned} L_i &= 2\,\tfrac{u_i-u_{i-1}}{h_i} - (1+\alpha_i)q_{i-1} - \beta_i q_i, \\ U_i &= 2\,\tfrac{u_i-u_{i-1}}{h_i} - \alpha_i q_{i-1} - (1+\beta_i)q_i \\ \Lambda_i &= \tfrac{\alpha_i^2}{\kappa_{i-1}} + \tfrac{\beta_i^2}{\kappa_i} + \tfrac{4}{h_i^2}\,(\tfrac{1}{\delta_{i-1}} + \tfrac{1}{\delta_i}), \\ \Gamma_i &= \tfrac{\alpha_i}{\kappa_{i-1}} + \tfrac{\beta_i}{\kappa_i} + \Lambda_i, \\ \Gamma_{1i} &= \tfrac{1+2\alpha_i}{\kappa_{i-1}} + \Lambda_i, \\ \Gamma_{2i} &= \tfrac{1+2\beta_i}{\kappa_i} + \Lambda_i, \end{aligned} \tag{5.18}$$

$$\begin{aligned} \S_i(\varrho,\xi,\sigma,\eta) &= \tfrac{\alpha_i}{\kappa_{i-1}}\xi + \tfrac{\beta_i}{\kappa_i}\eta + \tfrac{2}{h_i}(\tfrac{\varrho}{\delta_{i-1}} - \tfrac{\sigma}{\delta_i}), \\ X_i(\varrho,\xi,\sigma,\eta) &= 2L_i - \tfrac{\xi}{\kappa_{i-1}} - S_i(\varrho,\xi,\sigma,\eta), \\ Y_i(\varrho,\xi,\sigma,\eta) &= 2U_i - \tfrac{\eta}{\kappa_i} - S_i(\varrho,\xi,\sigma,\eta). \end{aligned}$$

Proposition 8 *In convex smoothing by the fit-and-modify method, the first order derivatives of the conjugates to the functions (4.7) and the sets (4.5) are*

$$\begin{aligned} \partial_1 H_i^*(\varrho,\xi,\sigma,\eta) &= u_{i-1} + \tfrac{\varrho}{2\delta_{i-1}} + \tfrac{\mu_i-\lambda_i}{h_i\delta_{i-1}}, \\ \partial_2 H_i^*(\varrho,\xi,\sigma,\eta) &= q_{i-1} + \tfrac{\xi}{2\kappa_{i-1}} - \tfrac{1+\alpha_i}{2\kappa_{i-1}}\,\lambda_i + \tfrac{\alpha_i}{2\kappa_{i-1}}\,\mu_i, \\ \partial_3 H_i^*(\varrho,\xi,\sigma,\eta) &= u_i + \tfrac{\sigma}{2\delta_i} + \tfrac{\lambda_i-\mu_i}{h_i\delta_i}, \\ \partial_4 H_i^*(\varrho,\xi,\sigma,\eta) &= q_i + \tfrac{\eta}{2\kappa_i} - \tfrac{\beta_i}{2\kappa_i}\lambda_i + \tfrac{1+\beta_i}{2\kappa_i}\,\mu_i, \end{aligned} \tag{5.19}$$

with the following functions $\lambda_i = \lambda_i(\varrho,\xi,\sigma,\eta)$ *and* $\mu_i = \mu_i(\varrho,\xi,\sigma,\eta)$,

$$\begin{aligned} &\lambda_i = 0,\ \mu_i = 0 \quad \text{for} \quad X_i \geq 0,\ Y_i \leq 0, \\ &\lambda_i = -X_i\Gamma_{1i},\ \mu_i = 0 \quad \text{for} \quad 0 \geq \Gamma_i X_i \geq \Gamma_{1i} Y_i, \\ &\lambda_i = 0,;\mu_i = -Y_i\Gamma_{2i} \quad \text{for} \quad 0 \leq \Gamma_i Y_i \leq \Gamma_{2i} X_i, \end{aligned} \tag{5.20}$$

$$\lambda_i = \frac{\Gamma_i Y_i - \Gamma_{2i} X_i}{\Gamma_{1i}\Gamma_{2i} - \Gamma_i \Gamma_i};$$

$$\mu_i = \frac{\Gamma_{1i} Y_i - \Gamma_i X_i}{\Gamma_{1i}\Gamma_{2i} - \Gamma_i \Gamma_i} \quad \text{for} \quad \Gamma_{1i} Y_i \geq \Gamma_i X_i\,, \;\; \Gamma_{2i} X_i \leq \Gamma_i Y_i\,.$$

Proof In view of the Lagrangian

$$\Phi = -\varrho f - \xi x - \sigma g - \eta y + \kappa_{i-1}(x - q_{i-1})^2 + \kappa_i (y - q_i)^2 + \delta_{i-1}(f - u_{i-1})^2 + \delta_i (g - u_i)^2 + \\ + \lambda((1+\alpha_i)x + \beta_i y - 2 \\ \tfrac{g-f}{h_i}) + \mu(2\,\tfrac{g-f}{h_i} - \alpha_i x - (1+\beta_i)\,y),$$

the equalities $\phi_f = 0$, $\phi_x = 0$, $\phi_g = 0$, $\phi_y = 0$ in the Kuhn-Tucker conditions lead to

$$f = u_{i-1} + \frac{\varrho}{2\delta_{i-1}} + \frac{\mu - \lambda}{h_i \delta_{i-1}}, \tag{5.19}$$

$$x = q_{i-1} + \tfrac{\xi}{2\kappa_{i-1}} - \tfrac{1+\alpha_i}{2\kappa_{i-1}}\,\lambda + \tfrac{\alpha_i}{2\kappa_{i-1}}\mu,$$

$$g = u_i + \tfrac{\sigma}{2\delta_i} + \tfrac{\lambda - \mu}{h_i \delta_i}\,, \;\; y = q_i + \tfrac{\eta}{2\kappa_i} - \tfrac{\beta_i}{2\kappa_i}\lambda + \tfrac{1+\beta_i}{2\kappa_i}\mu.$$

Thus we obtain straightforwardly

$$2\Phi_\lambda = -X_i - \Gamma_{1i}\lambda + \Gamma_i \mu, \;\; 2\Phi_\mu = Y_i + \Gamma_i \lambda - \Gamma_{2i}\mu\,. \tag{5.20}$$

In the four cases implied by the complementarity conditions (5.16), by means of (5.22) the functions $\lambda = \lambda_i$ and $\mu = \mu_i$ are obtained as given in (5.20). Likewise, the subdomains (5.20) are easily seen to be consequences of (5.16) and (5.22). In view of (5.21), the proof is complete if the formulae (5.17) are taken into account. ■

6 Computational Remarks

Computational Remarks
In constructing convex interpolants or approximants, here quadratic C^1 splines on refined grids are used. Thus, for data sets in strictly convex position, convex interpolation is always successful while in convex approximation this holds true without any assumptions. We are led to finite dimensional quadratic optimization problems. The programs are tridiagonally separable, and they can be effectively solved via dualization. The presented results suggest to proceed as follows.

Step 1: In order to make the interpolation problems solvable, choose the ratios of the additional nodes according to (3.9). In approximation, except feasibility there are no restrictions on the ratios.

Step 2: Solve the corresponding unconstrained dual programs (2.2) numerically.

Step 3: Determine the solutions of the primal programs (1.2) by means of the explicit return-formula (2.7). Compute the resulting convex splines (3.5), (3.6) and, if it is desired, plot them.

Step 4: Modify the weights in the Schoenberg functional ((4.2), or in (4.7), or in (3.15) if the visual behaviour of the obtained splines is to improve. Especially the choice of the weights κ, $\kappa_0, \ldots, \kappa_n$ is of importance. If these weights tend to zero the spline solutions interpolate more and more, while for large weights a straight-line is approximated.

In Step 2 computer tests with various types of tridiagonally separable programs have shown that Newton's method combined with the steepest descent method for supplying sufficiently good initial vectors is effective. In both methods the special structure can be taken into account easily. Of course, there are situations in which these methods fail. This happens, e.g., for very small weights κ or $\kappa_0, \ldots, \kappa_n$. Further, numerical problems arise in data sets with nearly equality on several places in the convexity assumption (3.3). However, we have not met with difficulties caused by the lack of differentiability in the Fenchel conjugates. Usually Newton's method terminates after a very moderate number of steps provided the initial vectors are suitable.

References

[1].- W. Burmeister, W. Heß, and J.W. Schmidt. *Convex spline interpolants with minimal curvature*, Computing, 35, 1985, 219-229.

[2].- P. Costantini. *On monotone and convex spline interpolation*, Math. Comp., 46, 1986, 203-214.

[3].- S. Dietze. *On partially separable optimization problems and duality*, Preprint-Reihe Mathematik 1, Universitt Greifswald, 1994, 9-12.

[4].- S. Dietze and J. W. Schmidt. *Determination of shape preserving spline interpolants with minimal curvature via dual programs*, J. Approx. Theory, 52(1), 1988, 43-57.

[5].- J. E. Falk. *Lagrange multipliers and nonlinear programming*, J. Math. Anal. Appl., 19, 1967, 141-159.

[6].- J. A. Holladay. *Smoothest curve approximation*, Math. Tables Aids Comput., 11, 1957, 233-243.

[7].- R. T. Rockafellar. *Convex Analysis*, Princeton Univ. Press, 1970.
[8].- J. W. Schmidt. *Dual algorithms for solving convex partially separable optimization problems*, Jber. d. Dt. Math.-Verein., 94, 1992, 40-62.
[9].- J. W. Schmidt *Staircase algorithm and convex spline interpolation*, Z. Angew. Math. Mech., 75, 1995, S661-S662.
[10].- J. W. Schmidt and W. Heß. *Schwach verkoppelte Ungleichungssysteme und konvexe Spline-Interpolation*, Elem. Math., 39, 1984, 85-95.
[11].- J.W. Schmidt. *On tridiagonal linear complementarity problems*, Numer. Math., 51, 1987, 11-21.
[12].- J.W. Schmidt and S. Dietze. *Unconstrained duals to partially separable constrained programs*, Math. Programming, 56, 1992, 337-341.
[13].- I.J. Schoenberg. *Spline functions and the problem of graduation*, Proc. Nat. Acad. Sci. U.S.A., 52, 1964, 947-950.
[14].- L. L. Schumaker. *On shape-preserving quadratic spline interpolation*, SIAM J. Numer. Anal., 20, 1993, 854-864.