International Journal of Shape Modeling, submitted

COMPUTATIONAL METHODS FOR DISCRETE PARAMETRIC ℓ_1 AND ℓ_∞ CURVE FITTING

BERT JÜTTLER
Dept. of Mathematics, University of Technology, Darmstadt
Schloßgartenstr. 7, 64289 Darmstadt, Germany
juettler@mathematik.tu-darmstadt.de

Received January 21, 1998
Revised June 2, 1998
Communicated by (Name of Editor)

The paper is devoted to ℓ_1 and ℓ_∞ approximation with parametric spline curves. We discuss the questions of existence and uniqueness of solutions. With the help of a suitable linearization of the Euclidean norm, we derive a method for computing the approximating spline curves. The method uses linear and quadratic programming in order to find the solution.

Keywords: Curve fitting, parametric spline curves, ℓ_1– and ℓ_∞–approximation.

1. Introduction

The construction of curves (and surfaces) from scattered data is one of the fundamental tasks of Computer Aided Geometric Design. It has attracted a great deal of research, see [1, Section 4.4] for references. A typical strategy consists of two steps. Assume, that a sequence of points $(\mathbf{p}_j)_{j=0,\ldots,n}$ in the plane $\mathbb{R}^2$ is given. Firstly one associates a sequence of parameters t_j with the data. Secondly, a suitable approximation scheme is applied to the coordinates of the data in order to find the approximating curve $\mathbf{x}(t)$, e.g., a B–spline curve.

The majority of the spline fitting procedures described in the literature relies on the ℓ_2–norm of the error vectors,

$$\sqrt{\sum_{j=0}^{n} \|\mathbf{x}(t_j) - \mathbf{p}_j\|^2}, \tag{1}$$

where $\|.\|$ is the Euclidean norm. (See Section 3 for more comments on the notation.) As an important advantage of this approach, the error function separates into independent terms for the two coordinates of the data. In addition, the solution can easily computed by solving a system of linear equations. As a disadvantage, the geometric meaning of the objective function (1) is somewhat complicated. The contribution of a point to (1) increases quadratically with its distance from the curve. That is, remote points (sometimes also called 'outliers'; cf. Section 4.5) have a much bigger influence to the resulting curve than the closer ones.

In some applications it may be more appropriate to use other kinds of error functions, in order to obtain better shapes. For instance, as an obvious choice, one may wish to minimize the *maximum distance* of the points $\mathbf{x}(t_j)$ from the data. This will produce the curve which is as close as possible to the given data. However, as a well–known fact, the solution to this problem tends to oscillations. Another suitable objective function could be the *overall length* of the error vectors $\mathbf{x}(t_j) - \mathbf{p}_j$. This is mainly advantageous for fitting so–called 'uncertain data' (as sometimes generated by optical scanning devices) which contain 'outliers', see [2]. These two possibilities lead to the discrete parametric ℓ_1 (overall length) and ℓ_∞ (maximum distance) approximation problems; they can be formulated with the help of the ℓ_1 and the ℓ_∞ norm of the sequence $(\|\mathbf{x}(t_j) - \mathbf{p}_j\|)_{j=0,\dots,n}$.

In the case of approximation with spline *functions*, both tasks are well–studied problems in approximation theory. Section 2 provides an outline of these results. As one of the basic observations, the ℓ_1 and ℓ_∞ approximation of scattered data with spline functions can be formulated as linear programming problems, see the note by Barrodale and Young [3], cf. [4]. A more recent contribution to this subject is the conference article by Heidrich et al. [2].

The present paper is devoted to both theoretical and computational aspects of ℓ_1 and ℓ_∞ approximation with *parametric* spline curves. Firstly, in Section 2, we summarize some of the results on ℓ_1, ℓ_2, ℓ_∞ data fitting with spline *functions* which are available in the literature. In Section 3 we formulate the approximation problems for *parametric* curves. We examine the questions of existence and uniqueness of solutions. It is shown that both problems do not have unique solutions in general. In order to guarantee uniqueness, we propose a modification of the original problems. The final section is devoted to computational aspects of the approximation problems. Based on a suitable linearization of the Euclidean norm we develop a method for parametric ℓ_1 and ℓ_∞ approximation. Finally, the scheme is illustrated by an example. We conclude the paper with a brief comparison of the different objective functions with respect to the shapes of the solutions.

2. Fitting Discrete Data with Spline Functions

A set of $n+1$ data $(t_j, p_j)_{j=0,\dots,n}$ with monotonically increasing abscissas, $t_j < t_{j+1}$, is assumed to be given. We consider a spline function in B–spline representation,

$$f(t) = \sum_{i=0}^{m} N_i^d(t)\, c_i, \quad t \in [t_0, t_n], \tag{2}$$

of degree d with the $m+1$ B-spline coefficients $c_i \in \mathbb{R}$, $m \leq n$. The B–spline functions $N_i^d(t)$ are defined with respect to a certain knot vector. For more information on spline functions and their B–spline representation the reader is referred to one of the various textbooks on this subject, e.g. [1, 5, 6]. The B–spline coefficients $c_i \in \mathbb{R}$ are to be chosen so that the spline function approximates the given data.

Of course, firstly one has to generate suitable knots for the spline function. An adaptive algorithm is described in the textbook by Dierckx [7].

As the most popular choice, the spline coefficients are often computed by minimizing the ℓ_2 norm of the error vector,

$$(\text{FA}_2 :) \quad \sqrt{\sum_{j=0}^{n}(f(t_j) - p_j)^2} = \|(f(t_j) - p_j)_{j=0,\dots,n}\|_2 \to \text{Min.} \tag{3}$$

This will be called the functional ℓ_2 approximation problem FA_2. As an important advantage of this error function, the spline coefficients c_i can be computed by solving a banded $(m+1) \times (m+1)$ system of linear equations. This system is formed by the so–called normal equations of the error function, see [8]. A unique solution can be shown to exist if and only if a subset of the abscissas t_j satisfies the Schoenberg–Whitney conditions[a] [8]. That is, if only this subset of the data would be given, then a unique interpolating spline function would exist.

In certain applications it may be more appropriate to use other kinds of error functions. For instance, for the fitting of uncertain data, the ℓ_1 norm of the error vector

$$(\text{FA}_1 :) \quad \sum_{j=0}^{n} |f(t_j) - p_j| = \|(f(t_j) - p_j)_{j=0,\dots,n}\|_1 \to \text{Min} \tag{4}$$

gives better results, as the approximating spline function simply ignores so–called 'outliers' [2]. This approximation will be called the functional ℓ_1 approximation problem FA_1. On the other hand, in order to minimize the global error, one may prefer to minimize the ℓ_∞ norm of the error vector,

$$(\text{FA}_\infty :) \quad \max_{j=0,\dots,n} |f(t_j) - p_j| = \|(f(t_j) - p_j)_{j=0,\dots,n}\|_\infty \to \text{Min.} \tag{5}$$

These leads to the functional ℓ_∞ approximation problem FA_∞.

As observed by Barrodale and Young in 1966, both approximation tasks can easily be formulated as linear programming (LP) problems, see [3] and the recent article [2]. The problem FA_1 is equivalent to the LP problem

$$(\text{FA}_1 :) \quad \begin{cases} \displaystyle\sum_{j=0}^{n} \epsilon_j \to \text{Min} \\ \text{subject to} \quad -\epsilon_j \le \displaystyle\sum_{i=0}^{m} N_i^d(t_j)\, c_i - p_j \le \epsilon_j \quad (j = 0, \dots, n) \end{cases} \tag{6}$$

with the $m+n+2$ variables $(\epsilon_j)_{j=0,\dots,n}$ and $(c_i)_{i=0,\dots,m}$. Similarly, the problem FA_∞ is equivalent to the LP problem

$$(\text{FA}_\infty :) \quad \begin{cases} \epsilon \to \text{Min} \\ \text{subject to} \quad -\epsilon \le \displaystyle\sum_{i=0}^{m} N_i^d(t_j)\, c_i - p_j \le \epsilon \quad (j = 0, \dots, n) \end{cases} \tag{7}$$

[a]The $m+1$ abscissas $t_{j(0)}, t_{j(1)}, \dots, t_{j(m)}$ are said to fulfill the Schoenberg–Whitney conditions, if $N_i^d(t_{j(i)}) > 0$ holds for the B–spline functions , $i = 0, \dots, m$. See [6, §4.8] for more information.

with the $m+2$ variables ϵ and $(c_i)_{i=0,\ldots,m}$. A number of algorithms for linear programming are available, see e.g. [9]. A computationally efficient algorithm for solving the ℓ_1 approximation problem is described in [4]. The LP formulation of the ℓ_∞ problem has been used by Esch and Eastman [10] in order to derive an algorithm for continuous best approximation. For more information on functional approximation schemes we refer to the survey by Nürnberger [11].

3. Fitting Discrete Data with Parametric Spline Curves

This article is devoted to curve fitting with planar *parametric* spline curves. A sequence of data $(\mathbf{p}_j)_{j=0,\ldots,n}$ in $\mathbb{R}^2$ with an associated sequence of monotonically increasing parameter values $(t_j)_{j=0,\ldots,n}$ is assumed to be given. There exist several algorithms for estimating the parameters from the data, see [1, Section 4.4.1]. We consider a polynomial parametric spline curve of degree d,

$$\mathbf{x}(t) = \sum_{i=0}^{m} N_i^d(t)\,\mathbf{d}_i, \quad t \in [t_0, t_n], \tag{8}$$

with the B-spline control points $\mathbf{d}_i \in \mathbb{R}^2$, see [1]. The control points $\mathbf{d}_i \in \mathbb{R}^2$ are to be chosen so that the spline curve approximates the given data. As in the previous section we assume that the knots of the approximating spline curve are already known.

The B–spline control points $\mathbf{d}_i$ are to be chosen so that the spline curve (8) approximates the given data. Similar to the functional case, it is a very popular approach [1] to compute the unknown spline coefficients (control points) by minimizing the ℓ_2 norm of the error vector,

$$(\mathrm{PA}_2:) \quad \sqrt{\sum_{j=0}^{n} \|\mathbf{x}(t_j) - \mathbf{p}_j\|^2} = \big\| \|\mathbf{x}(t_j) - \mathbf{p}_j\|_{j=0,\ldots,n} \big\|_2 \to \mathrm{Min}, \tag{9}$$

with the Euclidean norm $\|\mathbf{y}\| = \sqrt{y_1^2 + y_2^2}$. This will be called the parametric ℓ_2 approximation problem PA_2. It splits into two separate ℓ_2 problems of the type FA_2 for the coordinates of the curve. As a consequence, the control points $(\mathbf{d}_i)_{i=0,\ldots,m}$ can easily be computed by solving a banded $(m+1) \times (m+1)$ system of linear equations with two right–hand sides. A unique solution of this system can be shown to exist if and only if a subset of the abscissas $(t_j)_{j=0,\ldots,n}$ satisfies the Schoenberg–Whitney conditions, see [8].

Throughout this paper we will assume that the parameter values $(t_j)_{j=0,\ldots,n}$ are kept constant, they are not subject to optimization. Consequently, the curve (8) approximates both the given data *and* their parameterization. Hence the error vectors $\mathbf{x}(t_j) - \mathbf{p}_j$ are generally not perpendicular to the curve tangent at $t = t_j$, cf. Figure 3 (b). In order to improve the result of the approximation, by making the error vectors 'more orthogonal' to the curve, one may modify the parameter values with the help of the method of *parameter correction.* For any details the reader is referred to [1, Section 4.4.3].

In the sequel we need to distinguish between various norms. We denote with

$$\| (y_j)_{j=0,\ldots,n} \|_p \quad \text{for} \quad p \in \{1, 2, \infty\} \tag{10}$$

the ℓ_p norm of a vector from $\mathbb{R}^{n+1}$. On the other hand, we will use the abbreviation $\|\mathbf{y}\|$ for the Euclidean norm of a vector $\mathbf{y} = (\, y_1 \; y_2 \,)^\top \in \mathbb{R}^2$.

The remainder of this section is organized as follows. Firstly we introduce the parametric analogues of the ℓ_1 and ℓ_∞ approximation problems FA_1, FA_∞. Their solutions are shown to form a convex compact set. However, generally no *uniqueness* of the solution can be expected. Therefore, we propose to modify the original problems in order to get a unique solution.

3.1. *Parametric ℓ_1 and ℓ_∞ Curve Fitting*

Whereas the ℓ_2 approximation of scattered data is particularly easy to compute, in certain applications it may be advantageous to use other kinds of error functions. For example, if one wants to fit a curve to uncertain data it is more appropriate to compute the control points by minimizing the ℓ_1 norm of the error vector,

$$(\mathrm{PA}_1 :) \quad \sum_{j=0}^{n} \|\mathbf{x}(t_j) - \mathbf{p}_j\| = \left\| \|\mathbf{x}(t_j) - \mathbf{p}_j\|_{j=0,\ldots,n} \right\|_1 \to \text{Min.} \tag{11}$$

This task will be called the parametric ℓ_1 approximation problem PA_1. On the other hand, if the global error is to be minimized, one will prefer for compute the control points by minimizing the ℓ_∞ norm of the error vector,

$$(\mathrm{PA}_\infty :) \quad \max_{j=0,\ldots,n} \|\mathbf{x}(t_j) - \mathbf{p}_j\| = \left\| \|\mathbf{x}(t_j) - \mathbf{p}_j\|_{j=0,\ldots,n} \right\|_\infty \to \text{Min}, \tag{12}$$

which leads to the parametric ℓ_∞ approximation problem PA_∞. Unlike the ℓ_2 problem PA_2, however, the parametric ℓ_1 and ℓ_∞ approximation problems do *not* split into separate problems for the coordinates. Thus, in the parametric case, the exact solution of PA_1 and PA_∞ cannot be found by linear programming. In Section 4 we describe a linearization technique which leads to an approximate solution.

With each spline curve (8) one may associate the point $\mathbf{D} = (\, \mathbf{d}_0, \mathbf{d}_1, \ldots, \mathbf{d}_m \,) \in \mathbb{R}^{2m+2}$, simply by collecting the coordinates of the control points.

Proposition 1. If a subset of the parameters $(t_j)_{j=0,\ldots,n}$ satisfies the Schoenberg–Whitney conditions, then the solutions $\mathbf{D}$ to each of the approximation problems PA_1 and PA_∞ form a compact convex subset of $\mathbb{R}^{2m+2}$.

Proof. 1.) Convexity of the error functions. Consider two B–spline curves $\mathbf{x}(t)$ and $\mathbf{x}^*(t)$ (see (8)) with control points $\mathbf{D}$ and $\mathbf{D}^*$. Then we have for $0 \le s \le 1$

$$\begin{aligned} &\left\| \|\, (1-s)\, \mathbf{x}(t_j) + s\, \mathbf{x}^*(t_j) - \mathbf{p}_j \|_{j=0,\ldots,n} \right\|_p \\ &\le \left\| (1-s)\, \left(\|\mathbf{x}(t_j) - \mathbf{p}_j\|_{j=0,\ldots,n} \right) + s\, \left(\|\mathbf{x}^*(t_j) - \mathbf{p}_j\|_{j=0,\ldots,n} \right) \right\|_p \\ &\le (1-s)\, \left\| \|\mathbf{x}(t_j) - \mathbf{p}_j\|_{j=0,\ldots,n} \right\|_p + s\, \left\| \|\mathbf{x}^*(t_j) - \mathbf{p}_j\|_{j=0,\ldots,n} \right\|_p . \end{aligned} \tag{13}$$

Thus, both error functions $\|\ldots\|_p$, $p \in \{1, \infty\}$, are convex continuous functions of the control points.

2.) Compactness. Let $(t_{j(i)})_{i=0,\ldots,m}$ be a sub–sequence of the associated parameters $(t_j)_{j=0,\ldots,n}$ which satisfies the Schoenberg–Whitney conditions. Consequently, the matrix $N = (N_i^d(t_{j(k)}))_{i,k=0,\ldots,m}$ is invertible. Hence one may uniquely describe the spline curve (8) by its points $\mathbf{x}_i = \mathbf{x}(t_{j(i)})$, $i = 0, \ldots, m$, as

$$\mathbf{D} = N^{-1} \cdot \mathbf{X} \quad \text{with} \quad \mathbf{X} = (\mathbf{x}_i)_{i=0,\ldots,m} \tag{14}$$

holds. Let B be the value of the error function $\|\ldots\|_1$ or $\|\ldots\|_\infty$ for an arbitrary but fixed spline curve $\mathbf{x}^*$. Hence, the solutions to the problems PA_1, PA_∞, satisfy

$$B \geq \left\| \, \| \, \mathbf{x}(t_j) - \mathbf{p}_j \, \|_{j=0,\ldots,n} \right\|_p \geq \left\| \, \| \, \mathbf{x}_i - \mathbf{p}_{j(i)} \, \|_{i=0,\ldots,m} \right\|_p . \tag{15}$$

That is, the points $\mathbf{x}_i$ of the solutions to PA_1, PA_∞ are contained within the bounded set

$$\Omega = \{\mathbf{X} \mid \left\| \, \| \, \mathbf{x}_i - \mathbf{p}_{j(i)} \, \|_{i=0,\ldots,m} \right\|_p \leq B\}. \tag{16}$$

Owing to $\mathbf{D} = N^{-1} \cdot \mathbf{X}$, the control points of the solutions to PA_1, PA_∞ are contained within the bounded set $N^{-1} \cdot \Omega$. Thus, both error functions have their global minimum within the compact set $N^{-1} \cdot \Omega$, as they are convex and continuous. Therefore, the set of optimal points $\mathbf{D}$ is convex and compact. □

Remarks. 1.) Generally, the approximation problems PA_1 and PA_∞ do not have a unique solution. Two examples for this fact are shown in Figure 1. Both examples are B–spline curves of degree $d = 1$, i.e., polygonal lines. Whereas the boxes

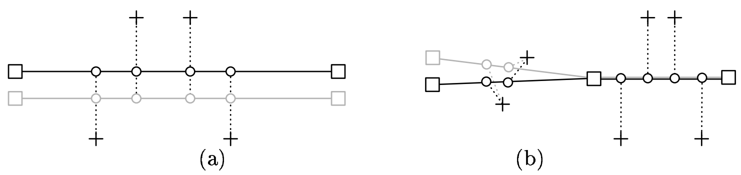

Fig. 1. Examples for non–unique solutions of the problems PA_1 (a) and PA_∞ (b)

represent their control points, the crosses and the circles mark the data $\mathbf{p}_j$ and the corresponding points $\mathbf{x}(t_j)$ on the curves, respectively. The Figures 1a and b show two solutions to the approximation problems PA_1 and PA_∞ for certain data. The solutions have been drawn as a black and as a grey curve. For many configurations of data $(\mathbf{p}_j)_{j=0,\ldots,n}$ and associated parameters $(t_j)_{j=0,\ldots,n}$, the problems PA_1 and PA_∞ can be expected to have unique solutions. If the data contains 'outliers', however, then the value of the objective function of PA_∞ (maximum distance) is only determined by the spline curve in the region of the 'outliers'. The remainder of the curve can then vary within a certain region, cf. Figure 1b.

2.) Clearly, computing the inverse N^{-1} might easily lead to numerical problems.

This inverse, however, is only required theoretically, in order to prove the compactness of the set of solutions. The solutions can be computed via linear and quadratic programming as described in Section 4.

3.2. *Modified Parametric ℓ_1 and ℓ_∞ Curve Fitting*

In order to guarantee the uniqueness of the solution, we modify the parametric ℓ_1 and ℓ_∞ approximation problems. The control points of the approximating spline curve (8) are found from

$$\begin{aligned} (\mathrm{MPA}_p :) \quad & \big\| \, \| \mathbf{x}(t_j) - \mathbf{p}_j \|_{j=0,\dots,n} \big\|_2 \to \mathrm{Min} \\ & \text{subject to} \quad \big\| \, \| \mathbf{x}(t_j) - \mathbf{p}_j \|_{j=0,\dots,n} \big\|_p \le \epsilon \\ & \text{where } \epsilon = \min\{ \, \big\| \, \| \bar{\mathbf{x}}(t_j) - \mathbf{p}_j \|_{j=0,\dots,n} \big\|_p \mid \bar{\mathbf{D}} \in \mathbb{R}^{2m+2} \, \} \end{aligned} \tag{17}$$

($p \in \{1, \infty\}$). That is, among the solutions of the approximation problems PA_1 and PA_∞ we choose the one which has the minimum ℓ_2 error. Clearly, instead of the ℓ_2 error function one might also use another quadratic function of the control points to pick the unique solution, e.g., a linearized 'energy' functional, see [1, Section 3.6]. According to our numerical experiences, however, the choice of the quadratic function is not so important, as most data will lead to unique solutions to the problems PA_p.

Proposition 2. If a subset of the parameters $(t_j)_{j=0,\dots,n}$ satisfies the Schoenberg–Whitney conditions, then the modified approximation problems MPA_1 and MPA_∞ have unique solutions.

Proof. According to Proposition 1, the feasible regions of MPA_1 and MPA_∞ in the control point space space $\{ \mathbf{D} = (\mathbf{d}_i)_{i=0,\dots,m} \mid \mathbf{D} \in \mathbb{R}^{2m+2} \}$ form a convex compact set. On the other hand, as a subset of the parameters $(t_j)_{j=0,\dots,n}$ is assumed to satisfy the Schoenberg–Whitney conditions, the least–squares sum

$$\| \, \| \mathbf{x}(t_j) - \mathbf{p}_j \|_{j=0,\dots,n} \|_2 \tag{18}$$

can be shown to be a strongly convex function of the control points, see [8]. Hence, a unique solution to MPA_1 and MPA_∞ exists. □

Thus, the modified approximation problems MPA_1 and MPA_∞ are a well–defined formulation for the task of parametric ℓ_1 and ℓ_∞ curve fitting. In the remainder of this paper we discuss the computational aspects of these problems.

4. Computation of Parametric ℓ_1 and ℓ_∞ Approximants

In the functional case, the ℓ_1 and ℓ_∞ approximation problems could be formulated as linear programming problems. In addition, the analogues of the modified ℓ_1 and ℓ_∞ approximation could simply be formulated as quadratic programming problems (that is, a quadratic objective function is to be minimized subject to linear equality and inequality constraints). For both types of programming problems, a number of powerful solvers exists [9]. According to theoretical results, the exact solution can be found in finite time.

In the parametric case, however, the situation becomes more difficult. In the sequel we present a linearization technique which can be used in order to compute approximate solutions to both modified parametric approximation problems MPA_1 and MPA_∞.

4.1. *Linearized Euclidean Norm*

In order to compute an approximate solution to the approximation problems MPA_1 and MPA_∞ we introduce a linearization of the Euclidean norm. We choose $p+1$ angles $(\phi_k)_{k=0,\ldots,p}$ with $0 \le \phi_0 < \phi_1 < \ldots < \phi_p < 2\pi$. With these angles we associate the $p+1$ unit vectors

$$\vec{\mathbf{u}}_k = \begin{pmatrix} \cos\phi_k \\ \sin\phi_k \end{pmatrix}, \quad k = 0, \ldots, p. \tag{19}$$

We assume that the oriented angle between neighbouring unit vectors is always smaller than π, i.e. $\phi_k - \phi_{k-1} < \pi$ $(k = 1, \ldots, p)$ and $2\pi + \phi_0 - \phi_p < \pi$. The dot '$\cdot$' stands for the inner product of two vectors. Let

$$[\mathbf{y}] = \max_{k=0,\ldots,p} \left(\frac{\vec{\mathbf{u}}_k + \vec{\mathbf{u}}_{(k+1) \bmod p}}{1 + \vec{\mathbf{u}}_k \cdot \vec{\mathbf{u}}_{(k+1) \bmod p}} \cdot \mathbf{y} \right) \tag{20}$$

The function $[.] : \mathbb{R}^2 \to \mathbb{R} : \mathbf{y} \mapsto [\mathbf{y}]$ is a piecewise linear function of the components of $\mathbf{y}$. It approximates the Euclidean norm $\|.\|$.

Lemma 3. For all vectors $\mathbf{y} \in \mathbb{R}^2$, the linearized norm $[.]$ satisfies the inequality

$$C\,\|\mathbf{y}\| \le [\mathbf{y}] \le \|\mathbf{y}\| \tag{21}$$

with the constant $C = \frac{1}{2} \min_{k=0,\ldots,p} \|\vec{\mathbf{u}}_k + \vec{\mathbf{u}}_{(k+1) \bmod p}\|$.

Proof. Consider the points $\mathbf{y} \in \mathbb{R}^2$ with $[\mathbf{y}] = 1$. These points form a polygon with the vertices $(\vec{\mathbf{u}}_k)_{k=0,\ldots,p}$, see Figure 2. On the one hand, the vertices are points on the Euclidean unit circle. On the other hand, the maximum radius of the inscribed circles with center $\mathbf{O}$ equals C. □

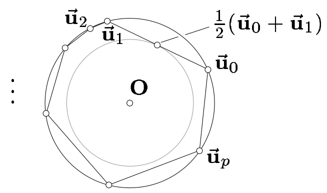

Fig. 2. Linearized Euclidean norm

Remark. By using the linearized norm (20), the points satisfying $[\mathbf{y}] = 1$ form an *inscribed* polygon to the unit circle. Alternatively, one might also use the linearized norm $[\mathbf{y}]^* = \max_{k=0,\ldots,p}(\mathbf{u}_k \cdot \mathbf{y})$ instead. Here, the points satisfying $[\mathbf{y}]^* = 1$ form a *circumscribed* polygon to the unit circle. Again, one may easily find an inequality similar to (21).

If we increase the number of angles $(\phi_k)_{k=0,\ldots,p}$ so that the angles between neighbouring unit vectors $\vec{\mathbf{u}}_k$ and $\vec{\mathbf{u}}_{(k+1)\text{mod } p}$ tend to zero, then the value of the linearized norm $[\mathbf{y}]$ converges to the Euclidean norm $\|\mathbf{y}\|$. The constant C measures the accuracy of the linearized Euclidean norm, see Lemma 3. We will quantify the accuracy by the value of C in %. The bigger the value of C is, the higher the accuracy gets.

4.2. *Linearized Parametric ℓ_1 and ℓ_∞ Curve Fitting*

With the help of the linearized Euclidean norm we introduce linearized versions of the approximation problems PA_1 and PA_∞. We simply replace the Euclidean norm $\|.\|$ with its linearized version $[.]$. This leads to the linearized parametric ℓ_1 approximation problem LPA_1,

$$(\text{LPA}_1 :) \quad \sum_{j=0}^{n} [\, \mathbf{x}(t_j) - \mathbf{p}_j \,] = \|\, [\, \mathbf{x}(t_j) - \mathbf{p}_j \,]_{j=0,\ldots,n} \,\|_1 \to \text{Min}, \qquad (22)$$

and to the linearized parametric ℓ_∞ approximation problem LPA_∞,

$$(\text{LPA}_\infty :) \quad \max_{j=0,\ldots,n} [\, \mathbf{x}(t_j) - \mathbf{p}_j \,] = \|\, [\, \mathbf{x}(t_j) - \mathbf{p}_j \,]_{j=0,\ldots,n} \,\|_\infty \to \text{Min}. \qquad (23)$$

Note that both tasks are *linear programming problems*, as the linearized norm $[.]$ is found as the maximum of a number of inner products. For instance, the first linearized problem may equivalently be formulated as

$$(\text{LPA}_\infty :) \quad \begin{cases} \displaystyle\sum_{j=0}^{n} \epsilon_j \to \text{Min} \\ \text{subject to } \left(\displaystyle\sum_{i=0}^{m} N_i^d(t_j)\mathbf{d}_i - \mathbf{p}_j \right) \cdot \dfrac{\vec{\mathbf{u}}_k + \vec{\mathbf{u}}_{(k+1)\text{mod } p}}{1 + \vec{\mathbf{u}}_k \cdot \vec{\mathbf{u}}_{(k+1)\text{mod } p}} \le \epsilon_j \\ \qquad (j = 0, \ldots, n;\ k = 0, \ldots, p). \end{cases} \qquad (24)$$

The properties of the solutions to the linearized problems are analogous to those of the original problems, cf. Proposition 1.

Proposition 4. If a subset of the parameters $(t_j)_{j=0,\ldots,n}$ satisfies the Schoenberg–Whitney conditions, then the solutions $\mathbf{D}$ of the linearized approximation problems LPA_1 and LPA_∞ form a compact convex subset of $\mathbb{R}^{2m+2}$.

As the proof is very similar to Proposition 1, it is omitted here.

In addition to the above proposition, the following fact can be shown. If we increase the number of angles $(\phi_k)_{k=0,\ldots,p}$ which are used for defining the linearized norm $[.]$ (see (20)) so that the angles between neighbouring unit vectors $\vec{\mathbf{u}}_k$ and

$\vec{\mathbf{u}}_{(k+1) \bmod p}$ tend to zero, then the solutions of the linearized approximation problems LPA_1 and LPA_∞ converge to the solutions of the original problems PA_1 and PA_∞. Once again, the linearized approximation problems LPA_1 and LPA_∞ do not have a unique solution in the general case. Similar to Figure 1 one may easily construct examples for non–uniqueness.

4.3. *Modified Linearized Parametric ℓ_1 and ℓ_∞ Curve Fitting*

In order to guarantee the uniqueness of the solution, we modify the parametric ℓ_1 and ℓ_∞ approximation problems. The control points of the approximating spline curve (8) are found from

$$
\begin{array}{ll}
(\mathrm{MLPA}_p :) & \| \, [\, \mathbf{x}(t_j) - \mathbf{p}_j \,]_{j=0,\ldots,n} \|_2 \to \mathrm{Min} \\
& \text{subject to} \quad \| \, [\, \mathbf{x}(t_j) - \mathbf{p}_j \,]_{j=0,\ldots,n} \|_p \le \epsilon \\
& \text{where} \quad \epsilon = \min\{ \, \| \, [\, \bar{\mathbf{x}}(t_j) - \mathbf{p}_j \,]_{j=0,\ldots,n} \|_p \mid \bar{\mathbf{D}} \in \mathbb{R}^{2m+2} \, \}
\end{array}
\tag{25}
$$

$(p \in \{1, \infty\})$. That is, among the solutions of the linearized approximation problems LPA_1 and LPA_∞ we choose the one which has the minimum ℓ_2 error.

Proposition 5. If a subset of the parameters $(t_j)_{j=0,\ldots,n}$ satisfies the Schoenberg–Whitney conditions, then the two modified approximation problems MLPA_1 and MLPA_∞ have unique solutions.

The proof is analogous to that of Proposition 2.

In addition to uniqueness of the solution, the following fact can be observed. If we increase the number of angles $(\phi_k)_{k=0,\ldots,p}$ which are used for defining the linearized norm [.] (see (20)) so that the angles between neighbouring unit vectors $\vec{\mathbf{u}}_k$ and $\vec{\mathbf{u}}_{(k+1) \bmod p}$ tend to zero, then the solutions of the linearized approximation problems LPA_1 and LPA_∞ converge to the solutions of the original problems PA_1 and PA_∞. On the one hand, the modified approximation problems MLPA_1 and MLPA_∞ are a well–defined linearized formulation for the task of parametric ℓ_1 and ℓ_∞ curve fitting. On the other hand, the solution is easy to compute. In the first step one has to solve a linear programming problem in order to find the value of the error function. In the second step, the final solution is found by solving a quadratic programming problem. That is, a quadratic objective function is minimized subject to linear inequality constraints. Both optimization problems are standard problems in optimization, see [9].

4.4. *Implementation*

Our implementation is based on the LOQO package [12] which is able to handle both linear and quadratic programming problems. The number of unknowns depends on the type of the problem MLPA_1 or MLPA_∞. For computing the linearized ℓ_1 approximation we have to use one error bound for each data point, as the objective function is the sum of all these error bounds. That is, we get optimization problems with $2\,(m+1) + (n+1)$ unknowns, where m and n are the number of control points and the number of data, respectively. For computing the linearized ℓ_∞ approximation, by contrast, we need to solve optimization problems with only $2\,(m+1)+1$ unknowns.

The number of linear inequalities depends on the number of the data and on the number p of the angles $(\phi_k)_{k=0,\dots,p}$ which are used in the linearization step, see (20). We get a system of $(n+1)(p+1)$ linear inequalities. The matrix of this system is sparse, as the B–spline functions are locally supported.

The accuracy of the solution depends on the constant on the left–hand side of the inequality from Lemma 3. In order to generate suitable angles $(\phi_k)_{k=0,\dots,p}$ for the linearized Euclidean norm we have tested two different strategies:

1.) If one chooses simply a *uniform distribution* of the angles, $\phi_k = \frac{2k}{p+1}\,\pi$, Lemma 3 gives the constant $C = \frac{1}{2}\sqrt{2 + 2\,\cos\frac{2\pi}{p+1}}$. That is, if we increase the number p of angles ϕ_k, then we increase the accuracy of the solutions to the parametric modified ℓ_1 and ℓ_∞ approximation problems. This, however, leads to large numbers of inequalities.

2.) In order to overcome this difficulty, one may use an *adaptive refinement* of the angles ϕ_k instead. One starts with a relatively small number of angles and computes an initial solution. With the help of this solution, the angle sequences (ϕ_k) are refined by inserting new angles where this is necessary. The refinement is based on the directions of the error vectors $\mathbf{x}(t_j) - \mathbf{p}_j$. Note that this step leads to individual sequences of angles for each one of the given points $\mathbf{p}_j$. Iterating this procedure a few times gives the final solution. As an advantage of this strategy, one may combine it with the method of *parameter correction* (see [1, Section 4.4.3]). Thus, one could simultaneously optimize both the parameterization of the data and the linearization of the Euclidean norm.

According to our numerical experiments, it is sufficient to use the first strategy for computing solutions with an accuracy (of the approximation to the Euclidean norm) of up to $C = 96\%$, cf. Lemma 3. In order to find highly accurate solutions, however, the second strategy should be preferred, as otherwise the number of inequalities gets too big.

In order to discuss the dependency of computing times and required computer memory from the number of data, we computed a series of approximations to data sets with different data sizes m, ranging from 100 to 1200. The number of control points was approximately $n = \frac{1}{3}m$. As indicated by these experiments, the computing time and the required amount of computing time for the optimization grow nearly linearly with the number of data. For instance, in order to fit a parametric spline curve with 200 control points to 600 data, the required optimization time was in the order of 35 seconds. Of course, computing the linearized ℓ_1 and ℓ_∞ approximation is more expensive than the traditional ℓ_2 fit. The asymptotic behaviour of the computing times, however, seems to be the same.

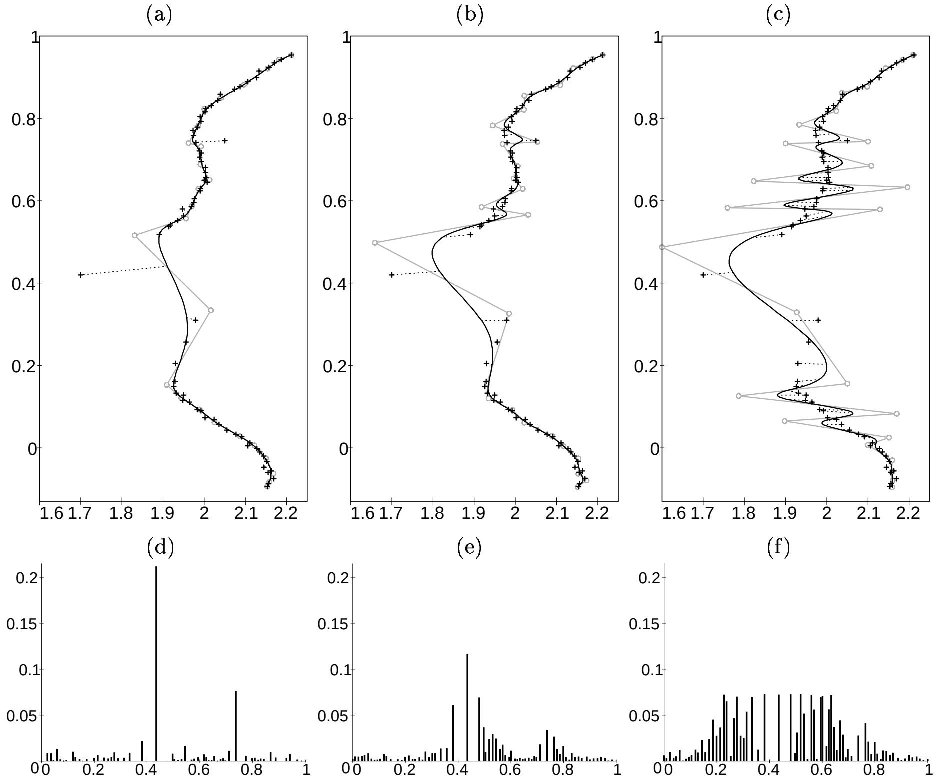

Fig. 3. Solutions of the approximation problems MLPA_1 (a), MLPA_∞ (c) and the ℓ_2 approximation (b). Length of the error vectors $(\mathbf{x}(t_j) - \mathbf{p}_j)_{j=0,\dots,n}$ (d,e,f).

4.5. *Examples*

As a first example we approximated 71 data with a centripetal parameterization (see [1]) with a cubic B–spline curve with 26 control points (23 segments). The data contains some 'outliers' and is irregularly distributed. The solutions to the approximation problems MLPA_1 and MLPA_∞ are shown in Figure 3a and c. The solutions have been computed with $p = 11$ and by choosing a uniform distribution of the angles ϕ_k. This leads to the accuracy $C = 96\%$ of the linearized norm. In order to compare the result, the ℓ_2 approximation of the data has been computed too, see Figure 3b. The first line of the Figure shows the data (marked by the crosses), the approximating spline curves with their control polygons, and the error vectors (dotted). In the second line, the lengths of the error vectors for the three approximating curves have been plotted. Table 1 shows the numerical values of the three error norms for the three spline curves.

Obviously, the approximating spline curves behave in very different ways. The ℓ_1 approximation (a,d) simply ignores 'outliers', i.e., data with relatively large distance from the remaining points. The overall length of the error vectors is minimized at the cost of the error at the 'outliers'. Consequently, the overall length of the error

Table 1. Comparison of ℓ_1–, ℓ_2, and ℓ_∞–approximation.

	ℓ_1 error	ℓ_2 error	ℓ_∞ error
MLPA_1	0.54	0.053	0.21
ℓ_2	0.76	0.030	0.12
MLPA_∞	1.84	0.092	0.073

vectors is relatively small. The maximum distance error, however, gets quite large. The ℓ_∞ approximation (c,f), by contrast, minimizes the maximum distance error. This is achieved at the cost of the overall error that is visualized by the area of the plot in Figure 3f. As a consequence, the resulting spline curve oscillates a lot. Finally, the ℓ_2 approximation (b,e) can be seen as a blend of the two extreme cases. It does not ignore the 'outliers', but they have less influence to the curve than in the ℓ_∞ case.

Another example is presented in Figure 4. The data have been sampled from a 'zig–zag' shape. Again, the figure shows only the ℓ_1, ℓ_2, and ℓ_∞ approximation. One may observe similar effects as in the first example. Note the differences at the corners of the data!

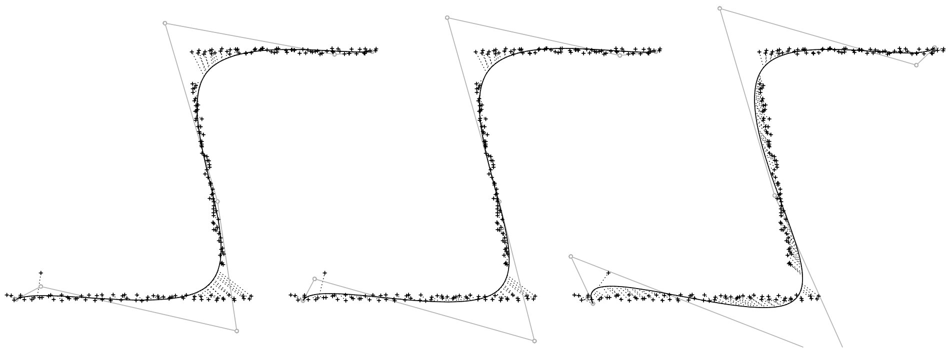

Fig. 4. Solutions of the approximation problems MLPA_1 (left), MLPA_∞ (right), and the ℓ_2–approximation (middle) for 'zig–zag' shaped data.

4.6. *Comparison of the error functions*

As observed in our experiments, if the data has no 'outliers', then in many cases there is only little difference between the ℓ_2 and the ℓ_∞ approximation. The computation of the ℓ_∞ approximation, however, is much more expensive. Whereas the solution of MLPA_∞ is found by solving a quadratic programming problem, the ℓ_2 approximation can simply be computed from a system of linear equations. Thus, we cannot recommend to use ℓ_∞ in practice. The ℓ_1 approximation, by contrast, is very well suited for data which contains 'outliers' ('uncertain data', see [2]). In many examples, the shape of the ℓ_1 approximation is much better than the shape of the ℓ_2 fit. Thus, for uncertain data, using the ℓ_1 approximation is a valuable alternative approach to data fitting, and the additional efforts may well be justified.

4.7. *Spatial Curves and Surfaces*

Finally, we give an outline of the generalization of the previously developed linearization method to the case of *spatial* curves and *surfaces.* In the planar case, we replace the unit *circle* of the exact Euclidean norm with an inscribed polygon. This becomes more difficult in the spatial case, as we have to deal with the unit *sphere.* Unlike the planar case, there is no canonical distribution of an arbitrary number of points on a sphere, and the construction of optimal distributions is a difficult task. See [13] for coordinates of putatively optimal distributions of n points on the sphere with $n \leq 130$. Summing up, the above approach to parametric ℓ_1 and ℓ_∞ can be generalized to spatial B–spline curves and to tensor–product B–spline surfaces, but its implementation gets more complicated.

Acknowledgement

The author wishes to thank the anonymous referees of this paper. Their careful reports have helped to improve the presentation of the material.

References

1. J. Hoschek and D. Lasser, *Fundamentals of Computer Aided Geometric Design* (AK Peters, Wellesley MA, 1993).
2. W. Heidrich, R. Bartels and G. Labahn, "Fitting Uncertain Data with NURBS", in *Curves and Surfaces with Applications in CAGD*, eds. A. Le Méhauté, C. Rabut and L.L. Schumaker (Vanderbilt University Press, Nashville, 1997), 177–184.
3. I. Barrodale and A. Young, "A note on numerical procedures for approximation by spline functions", *The Computer Journal* **9** (1966), 318-320.
4. I. Barrodale and F.D.K. Roberts, "An improved algorithm for discrete ℓ_1 linear approximation", *SIAM J. Numer. Anal.* **10** (1973), 839–848.
5. C. de Boor, *A Practical Guide to Splines* (Springer, New York, 1978).
6. L.L. Schumaker, *Spline Functions: Basic Theory* (Wiley, New York, 1981).
7. P. Dierckx, *Curve and Surface Fitting with Splines* (Clarendon Press, Oxford, 1993).
8. M.G. Cox, "Practical Spline Approximation", in *Topics in Numerical Analysis*, ed. P.R. Turner (Springer, Berlin, 1982), 79–111.
9. R. Fletcher, *Practical Methods of Optimization* (Wiley–Interscience, Chichester, 1990).
10. R.E. Esch and W.L. Eastman, "Computational Methods for Best Spline Function Approximation", *J. Approx. Theory* **2** (1969), 85–96.
11. G. Nürnberger, "Best Approximation by Spline Functions: Theory and Numerical Methods", in *Delay Equations, Approximation and Application, ISNM* **74**, eds. G. Meinardus and G. Nürnberger (Birkhäuser, Basel 1985), 180–212.
12. R. Vanderbei, *LOQO,* available at `ftp://ftp.zib.de/pub/opt-net/software/loqo/`.
13. R.H. Hardin, N.J.A. Sloane and W.D. Smith, "Spherical Coverings", available at `http://www.research.att.com/~njas/coverings/index.html`.