Computer-Aided Design 58 (2015) 179–188

Contents lists available at ScienceDirect

Computer-Aided Design

journal homepage: www.elsevier.com/locate/cad

Knot calculation for spline fitting via sparse optimization☆

Hongmei Kang, Falai Chen*, Yusheng Li, Jiansong Deng, Zhouwang Yang

School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, PR China

HIGHLIGHTS

- We reduce the computation time dramatically by solving convex optimization problem.
- We can simultaneously find a good combination of the knot number and knot locations.
- The algorithm has less knots with good fitting performance compared to other methods.
- We can recover the ground truth knots when data is sampled enough from a B-spline.

ARTICLE INFO

Keywords:
Spline fitting
Knot calculation
Sparse optimization

ABSTRACT

Curve fitting with splines is a fundamental problem in computer-aided design and engineering. However, how to choose the number of knots and how to place the knots in spline fitting remain a difficult issue. This paper presents a framework for computing knots (including the number and positions) in curve fitting based on a sparse optimization model. The framework consists of two steps: first, from a dense initial knot vector, a set of active knots is selected at which certain order derivative of the spline is discontinuous by solving a sparse optimization problem; second, we further remove redundant knots and adjust the positions of active knots to obtain the final knot vector. Our experiments show that the approximation spline curve obtained by our approach has less number of knots compared to existing methods. Particularly, when the data points are sampled dense enough from a spline, our algorithm can recover the ground truth knot vector and reproduce the spline.

1. Introduction

Curve fitting with splines is a traditional and fundamental problem in many engineering practices. In Computer Aided Design (CAD) and Geometric Modeling, curves are fitted with splines to reconstruct geometric models from measurement data [1–4]. In signal processing and image processing, splines are often adopted to process noisy signals or to approximate complicated functions [5,6].

The intuitive idea of curve fitting with splines is to formulate it as a least-square problem when knots are fixed. However, the fitting result is not always satisfactory. Actually, it has long been known that freeing knots in fitting improves the result dramatically [7–10]. But spline fitting with free knots is still a challenging problem. The reasons are as follows. First, analytic expressions for optimal knot locations, or even for general characteristics of optimal knot distributions, are not easy to derive [11]. Second, the unknown number and position of knots result in a large and nonlinear optimization problem, which is computationally very difficult.

In the literature many methods have been proposed to optimize knots with a given number of knots. The problem of knot placement is formulated as a nonlinear optimization problem with the constraint that knots should form a nondecreasing sequence. The first type of techniques transforms the constrained optimization problem into an unconstrained problem, then local gradient-based method or Gauss–Newton method are employed for minimization [11–13]. However, local optimization methods require a good initial guess and cannot guarantee global optimality. The second type of techniques applies global optimization to avoid the drawbacks of local methods, but it is computationally more expensive [14–16]. There are also some works which utilize the underlying feature information of the data to select knots, instead of solving a nonlinear optimization problem [1,17,18]. However, in such methods the number of knots is determined beforehand and the results are sensitive to measurement noises.

Another approach for knot calculation is based on knot-removal strategy which is to reduce the number of knots of a given spline

☆ This paper has been recommended for acceptance by Vadim Shapiro.

* Corresponding author.
E-mail addresses: khmkangkang@gmail.com, chenfl@ustc.edu.cn (F. Chen).

http://dx.doi.org/10.1016/j.cad.2014.08.022

by perturbing the spline within a given tolerance [19,20]. The main idea of the technique is to remove interior knots according to assigned weights. For data approximation, a piecewise linear approximation of the data is computed, then knot removal strategy is performed on the linear approximation, and finally the data is approximated by a smooth spline with computed knots.

A new development in recent years for knot calculation is based on sparse optimization [21,22]. Sparsity is the core of compressed sensing which is widely used in computer vision and signal processing [23–27]. Sparsity means that a signal can be represented in a linear combination of some bases or dictionaries such that most of the combination coefficients are zero. In [21], the authors formulated the spline fitting problem as a convex optimization problem, where the l_1 norm of jump of third order derivatives of C^2 cubic splines is minimized; while in [22], the authors first selected a subset of basis functions from the pre-specified multi-resolution basis set using the statistical variable selection method-Lasso, then identified a concise knot vector that is sufficient to characterize the vector space spanned by the selected basis functions to fit the data. These two methods can compute the number and positions of the knots simultaneously, yet they still produce a lot of redundant knots.

Targeting on the limitations of existing methods, we propose a computationally efficient framework to calculate knots for splines fitting via sparse optimization. The framework is composed of two stages: firstly we solve a convex sparse optimization model starting from a dense initial knot vector. The output is those knots (which we call active knots) at which a certain order derivatives of the fitting spline is discontinuous. The idea to formulate the optimization model in this step is the same as that in [21] but with a distinct formulation. Secondly, we adjust the active knots in the first stage by certain rules to remove redundant knots. Furthermore, several theoretical results about the algorithm are established in this paper. In particular, when the data points are sampled dense enough from a spline, the knots of this spline can be recovered by the proposed framework in any given precision.

The remainder of the current paper is organized as follows. In Section 2, we review some preliminary knowledge about B-splines and least-square fitting with B-splines. In Section 3, a two-stage framework of curve fitting with B-splines is described. Some related theoretical results are also presented. In Section 4, we illustrate the effectiveness of the proposed method through numerical experiments and comparisons with existing methods. Finally, in Section 5, we conclude the paper with discussions on future research problems.

2. Preliminaries

We refer to the fundamental book [1] for a complete treatment of splines. Here we simply introduce the adopted notations which are needed for presenting our results.

2.1. B-splines

Let $\{c_i\}_{i=0}^n \in \mathbb{R}^d$ be $n+1$ control points, and $N_i^p(t)$ be the B-spline basis functions of degree p defined on a knot vector $U = \{t_0, t_1, \dots, t_{n+p+1}\}$ with $t_i \le t_{i+1}, i = 0, 1, \dots, n+p$, then a B-spline curve of degree p is defined by

$$c(t) = \sum_{i=0}^{n} c_i N_i^p(t), \tag{1}$$

where $N_i^p(t)$ is defined recursively as follows:

$$N_i^0(t) = \begin{cases} 1, & t \in [t_i, t_{i+1}) \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$$N_i^p(t) = \frac{t - t_i}{t_{i+p} - t_i} N_i^{p-1}(t) + \frac{t_{i+p+1} - t}{t_{i+p+1} - t_{i+1}} N_{i+1}^{p-1}(t), \quad p \ge 1. \tag{3}$$

When $d = 1$, $c(t)$ is called a B-spline function and the control points are called spline coefficients. In this paper, we only consider data fitting with B-spline functions.

U is often chosen as an open knot vector, namely boundary knots are set to $a = t_0 = t_1 = \cdots = t_p$, $t_{n+1} = \cdots = t_{n+p+1} = b$. The knots t_i, $i = p+1, \dots, n$ are called interior knots of U. The multiplicity of an interior knot t_i is denoted by m_i ($m_i \le p+1$). An interior knot t_i is called an *active knot* of $c(t)$ if the $(p+1-m_i)$th order derivative of $c(t)$ is discontinuous at t_i, otherwise it is called an *inactive knot* of $c(t)$. Fig. 1 shows a C^2 continuous cubic B-spline function and its first three derivatives. The spline has 9 interior knots (marked by crosses) and 3 active knots (marked by black crosses) where the third order derivatives are discontinuous.

The kth order derivative of $c(t)$ is a B-spline of degree $p-k$:

$$c^{(k)}(t) = \Pi_{i=1}^{k}(p+1-i) \sum_{i=k}^{n} c_i^{(k)} N_i^{p-k}(t), \tag{4}$$

with

$$c_i^{(k)} = \begin{cases} c_i, & \textit{if } k = 0, \\ \dfrac{c_i^{(k-1)} - c_{i-1}^{(k-1)}}{t_{i+p+1-k} - t_i}, & \textit{if } k > 0. \end{cases} \tag{5}$$

The Fourier transform of the jth basis function $N_j^p(t)$ is defined as

$$\widehat{N}_j^p(t) = \int_{-\infty}^{-\infty} N_j^p(t) e^{iwt} dt = \frac{(p+1)!}{(iw)^{p+1}} \sum_{k=j}^{p+1+j} \frac{e^{iwt_k}}{\theta'(t_k)}, \tag{6}$$

where $\theta(t) = \Pi_{k=j}^{p+1+j}(t - t_k)$, $w \in \mathbb{R}$ represents frequency. For uniformly distributed knots, the Fourier transform can be simplified as:

$$\widehat{N}_j^p(t) = \left(\frac{e^{i|\tau|w} - 1}{i|\tau|w} \right)^{p+1}, \tag{7}$$

where $|\tau|$ is defined as $|\tau| = \max_i (t_{i+1} - t_i)$.

2.2. Least squares approximation by splines

Given a set of data $\{P_i\}_{i=1}^N$, and corresponding parameter values $\{s_i\}_{i=1}^N$, the least square approximation with splines is defined

$$\min_{c(t)} \sum_{i=1}^{N} (c(s_i) - P_i)^2, \tag{8}$$

where $c(t)$ is a spline function defined by (1). When the knot vector U of the spline function $c(t)$ is fixed, problem (8) is reduced to

$$\min_{C \in \mathbb{R}^{n+1}} \|P - AC\|^2, \tag{9}$$

where $P = (P_1, \dots, P_N)^T$, $A = (a_{ij})_{N \times (n+1)}$ with $a_{ij} = N_j^p(s_i)$, and $C = (c_0, c_1, \dots, c_n)^T$ is the coefficient vector.

If A has rank $n+1$, then A^TA is nonsingular, thus the solution C of problem (9) is obtained uniquely by

$$A^TAC = A^TP. \tag{10}$$

The sufficient and necessary conditions for A to have rank $n+1$ are stated by Schoenberg and Whitney [28]. If A does not have full rank, the solution C is defined as the solution which minimizes $\|C\|_2$ among all the solutions of (10).

3. Knot calculation for spline fitting

In this section, we will present a two-stage framework of knot calculation for spline fitting in detail. We start with an outline of the algorithm. Then the sparse optimization model and knot adjustment strategy are described respectively.