GEOMETRIC INTERPOLATION OF DATA IN $\mathbb{R}^3$

Jernej Kozak
Faculty of Mathematics and Physics and IMFM
Jadranska 19, SI-1000 Ljubljana, Slovenia
jernej.kozak@fmf.uni-lj.si

Emil Žagar
Faculty of Mathematics and Physics and IMFM
Jadranska 19, SI-1000 Ljubljana, Slovenia
emil.zagar@fmf.uni-lj.si

Abstract In this paper, the problem of geometric interpolation of space data is considered. Cubic polynomial parametric curve is supposed to interpolate five points in three dimensional space. It is a case of a more general problem, i.e., the conjecture about the number of points in $\mathbb{R}^d$ which can be interpolated by parametric polynomial curve of degree n. The necessary and sufficient conditions are found which assure the existence and the uniqueness of the interpolating polynomial curve.

Keywords: Parametric curve, geometric interpolation

1. Introduction

Interpolation by parametric polynomial curves is an important approximation procedure in computer graphics, computer aided geometric design, computer aided modeling, mathematical modeling,.... The word geometric refers to the fact that the interpolating curve is not forced to pass the points at given parameter values but is allowed to take its "minimum norm shape". It is well known too, that this kind of interpolation can lead to interpolation schemes of high order accuracy. In [4], the authors conjectured that a parametric polynomial curve of degree n in $\mathbb{R}^d$ can, in general, interpolate

$$n + 1 + \left\lfloor \frac{n-1}{d-1} \right\rfloor$$

data points. Some results by means of the asymptotic analysis can be found in [2], [5] and [8], but there are only a few results on this conjecture which do not involve asymptotic analysis, e.g., [3], [6] and [7]. In this paper the conjecture is proved to be true in the simplest nontrivial space case. More precisely, the cubic polynomial curve is found, which interpolates five points in $\mathbb{R}^3$. It is clear that this can not be done in general. The necessary and sufficient conditions on data points, which ensure the existence of the unique interpolating polynomial curve are provided. These conditions are purely geometric and do not require any asymptotic approach.

The problem which was described above, can be formalized as follows. Suppose five points $\boldsymbol{T}_j \in \mathbb{R}^3$, $j = 0, 1, 2, 3, 4$, are given. It is assumed that $\boldsymbol{T}_j \neq \boldsymbol{T}_{j+1}$. Is there a unique regular cubic parametric polynomial curve $\boldsymbol{B}$, which satisfies the interpolating conditions

$$\boldsymbol{B}(t_j) = \boldsymbol{T}_j, \quad j = 0, 1, 2, 3, 4, \tag{1}$$

where $t_0 < t_1 < t_2 < t_3 < t_4$ are unknown parameter values? Clearly, t_0 and t_4 can, e.g., be chosen as $t_0 := 0$ and $t_4 := 1$, since one can always apply a linear reparametrization. Thus the only unknown parameters left are t_1, t_2 and t_3 which have to lie in a domain

$$\mathcal{D} := \{\boldsymbol{t} := (t_1, t_2, t_3);\ 0 =: t_0 < t_1 < t_2 < t_3 < t_4 := 1\}. \tag{2}$$

Recall that $\boldsymbol{B}$ is a vector polynomial in $\mathbb{R}^3$, and its coefficients are also unknown. But once t_j, $j = 1, 2, 3$, are determined, any classical interpolation scheme on arbitrary four points trivially produces coefficients of $\boldsymbol{B}$. Thus the main problem is how to determine the parameters t_j. Since the interpolating polynomial curve is cubic, the problem is clearly nonlinear and one can expect the system of nonlinear equations. One of the ways how to obtain it is described in the next section.

2. The system of nonlinear equations

A polynomial curve, which satisfies (1), is cubic, and any divided difference on five points maps it to zero, i.e.,

$$[t_0, t_1, t_2, t_3, t_4]\boldsymbol{B} = \boldsymbol{0}. \tag{3}$$

Since t_j are different, the equations (3) can also be written as

$$\sum_{j=0}^{4} \frac{1}{\dot{\omega}(t_j)} \boldsymbol{B}(t_j) = \boldsymbol{0}, \tag{4}$$

where

$$\omega(t) := \prod_{j=0}^{4}(t-t_j), \quad \dot{\omega} := \frac{d\omega}{dt}.$$

By (1), the equation (4) rewrites to

$$\sum_{j=0}^{4} \frac{1}{\dot{\omega}(t_j)} \boldsymbol{T}_j = \mathbf{0}, \tag{5}$$

i.e., the system of three nonlinear equations for t_1, t_2 and t_3. Furthermore,

$$[t_0, t_1, t_2, t_3, t_4]\boldsymbol{T}_k = \sum_{j=0}^{4} \frac{1}{\dot{\omega}(t_j)} \boldsymbol{T}_k = \mathbf{0}, \tag{6}$$

and one of the terms in (5) can be always canceled by subtracting (6) for any k. This leads to scalar equations for the unknown t_j. More precisely, if (6) is subtracted from (5) with $k = 4$, e.g., the equation

$$\sum_{j=0}^{3} \frac{1}{\dot{\omega}(t_j)} (\boldsymbol{T}_j - \boldsymbol{T}_4) = \mathbf{0} \tag{7}$$

is obtained. Cross multiplication of (7) by $(\boldsymbol{T}_3 - \boldsymbol{T}_4)$ and scalar product by $(\boldsymbol{T}_2 - \boldsymbol{T}_4)$ lead to

$$\begin{aligned} &\frac{1}{\dot{\omega}(t_0)}((\boldsymbol{T}_0 - \boldsymbol{T_4}) \times (\boldsymbol{T}_3 - \boldsymbol{T}_4)) \cdot (\boldsymbol{T}_2 - \boldsymbol{T}_4) \\ &+\frac{1}{\dot{\omega}(t_1)}((\boldsymbol{T}_1 - \boldsymbol{T_4}) \times (\boldsymbol{T}_3 - \boldsymbol{T}_4)) \cdot (\boldsymbol{T}_2 - \boldsymbol{T}_4) = \mathbf{0}. \end{aligned} \tag{8}$$

Since $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$, a simple manipulation by determinants simplifies (8) to

$$\frac{\dot{\omega}(t_0)}{\dot{\omega}(t_1)} + 1 + \frac{\det(\Delta\boldsymbol{T}_0, \Delta\boldsymbol{T}_2, \Delta\boldsymbol{T}_3)}{\det(\Delta\boldsymbol{T}_1, \Delta\boldsymbol{T}_2, \Delta\boldsymbol{T}_3)} = \mathbf{0},$$

where $\Delta\boldsymbol{T}_j := \boldsymbol{T}_{j+1} - \boldsymbol{T}_j$. Two other nonlinear scalar equations can be derived in a similar way with different k applied in (7) , and we finally get the system

$$\begin{aligned} f_1(t_1, t_2, t_3; \alpha_1) &:= \frac{\dot{\omega}(t_0)}{\dot{\omega}(t_1)} + 1 + \alpha_1 = 0, \\ f_2(t_1, t_2, t_3; \alpha_2) &:= -\frac{\dot{\omega}(t_4)}{\dot{\omega}(t_0)} + \alpha_2 = 0, \\ f_1(t_1, t_2, t_3; \alpha_3) &:= \frac{\dot{\omega}(t_4)}{\dot{\omega}(t_3)} + 1 + \alpha_3 = 0, \end{aligned} \tag{9}$$

where

$$\begin{aligned} \alpha_1 &:= \frac{\det(\Delta \boldsymbol{T}_0, \Delta \boldsymbol{T}_2, \Delta \boldsymbol{T}_3)}{\det(\Delta \boldsymbol{T}_1, \Delta \boldsymbol{T}_2, \Delta \boldsymbol{T}_3)}, \\ \alpha_2 &= \frac{\det(\Delta \boldsymbol{T}_1, \Delta \boldsymbol{T}_2, \Delta \boldsymbol{T}_3)}{\det(\Delta \boldsymbol{T}_0, \Delta \boldsymbol{T}_1, \Delta \boldsymbol{T}_2)}, \\ \alpha_3 &= \frac{\det(\Delta \boldsymbol{T}_0, \Delta \boldsymbol{T}_1, \Delta \boldsymbol{T}_3)}{\det(\Delta \boldsymbol{T}_0, \Delta \boldsymbol{T}_1, \Delta \boldsymbol{T}_2)}. \end{aligned} \tag{10}$$

The system (9) can be shortly written as

$$\boldsymbol{F}(\boldsymbol{t}; \boldsymbol{\alpha}) := [f_1(t_1, t_2, t_3; \alpha_1), f_2(t_1, t_2, t_3; \alpha_2), f_3(t_1, t_2, t_3; \alpha_3)]^T = \mathbf{0},$$

where $\boldsymbol{\alpha} := [\alpha_1, \alpha_2, \alpha_3]^T$. The main theorem of this paper is now the following.

Theorem 1 *A cubic parametric curve through five points $\boldsymbol{T}_j \in \mathbb{R}^3$, $j = 0, 1, 2, 3, 4$, is uniquely determined if and only if the components of $\boldsymbol{\alpha}$, defined by (10), are all positive.*

3. The proof of the theorem

In this section the proof of the Theorem 1 will be given.

If the system (9) has a unique solution in $\mathcal{D}$ defined by (2), then a straightforward computation shows that

$$\frac{\dot\omega(t_0)}{\dot\omega(t_1)} < -1, \quad \frac{\dot\omega(t_4)}{\dot\omega(t_0)} > 0, \quad \text{and} \quad \frac{\dot\omega(t_4)}{\dot\omega(t_3)} < -1.$$

This implies that α_i, $i = 1, 2, 3$, must be positive and the first part of the theorem is proved.

The proof that the positivity of the components of $\boldsymbol{\alpha}$ is also sufficient condition, will be split into two main parts.

a) A unique solution of the system for a particular vector $\boldsymbol{\alpha}^*$ will be established.

b) The fact that a unique solution exists will be extended to all admissible vectors $\boldsymbol{\alpha}$ by the aim of the homotopy theory.

a) Consider a particular system (9) first, i.e.,

$$\boldsymbol{F}(\boldsymbol{t}; \boldsymbol{\alpha}^*) = \mathbf{0}, \quad \boldsymbol{\alpha}^* := [3, 1, 3]^T. \tag{11}$$

Its polynomial equivalent on $\mathcal{D}$ reads

$$\begin{aligned} p_1(t_1, t_2, t_3) &:= f_1(t_1, t_2, t_3; 3)\, \dot\omega(t_1) = 0, \\ p_2(t_1, t_2, t_3) &:= f_2(t_1, t_2, t_3; 1)\, \dot\omega(t_0) = 0, \\ p_3(t_1, t_2, t_3) &:= f_3(t_1, t_2, t_3; 3)\, \dot\omega(t_3) = 0. \end{aligned} \tag{12}$$

One of the possible approaches to such polynomial systems is to use resultants as a tool that brings the system to a higher degree single variable case. Let $\mathrm{Res}(p, q, x)$ denote the resultant of polynomials p, and q, with respect to the variable x. It is straightforward to compute

$$\mathrm{Res}(\mathrm{Res}(p_1, p_2, t_2), \mathrm{Res}(p_2, p_3, t_2), t_3) = 16\, t_1^{10}\, (1 - t_1)^{10}\, q(t_1),$$

where

$$q(t_1) := 1024\, t_1^6 - 3072\, t_1^5 + 5952\, t_1^4 - 6784\, t_1^3 + 4392\, t_1^2 - 1512\, t_1 + 189.$$

Since $t_1 \neq 0, 1$, the only candidates for the first component of the solution $\boldsymbol{t}$ are the six roots of the polynomial q, i.e.,

$$\frac{1}{4}\,,\ \frac{3}{4}\,,\ \frac{2 \pm i\sqrt{17 \pm 3\sqrt{21}}}{4}\,.$$

The second equation in (12) is obviously linear in t_2, and it is easy to deduce that only the root $t_1 = \dfrac{1}{4}$ produces the (unique) solution

$$\boldsymbol{t} = \left[\frac{1}{4},\ \frac{2}{4},\ \frac{3}{4}\right]^T \tag{13}$$

of the system (11) in $\mathcal{D}$.
b) In order to extend the fact from a) to the general $\boldsymbol{\alpha}$, consider the linear homotopy

$$\boldsymbol{H}(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda) := (1 - \lambda)\, \boldsymbol{F}(\boldsymbol{t}; \boldsymbol{\alpha}^*) + \lambda\, \boldsymbol{F}(\boldsymbol{t}; \boldsymbol{\alpha}) \tag{14}$$

A particular form of the Brouwer's degree of a differentiable map $\boldsymbol{G}$ reads

$$\mathrm{degree}(\boldsymbol{G}, \mathcal{D}) = \sum_{\boldsymbol{t} \in \mathcal{D},\, \boldsymbol{G}(\boldsymbol{t}) = \boldsymbol{0}} \mathrm{sign}(\det(J(\boldsymbol{G})(\boldsymbol{t}))), \tag{15}$$

where J is a Jacobian of $\boldsymbol{G}$ with respect to $\boldsymbol{t}$. It gives some information about the number of the zeros of $\boldsymbol{G}$ in $\mathcal{D}$. In particular, if

$$\mathrm{degree}(\boldsymbol{G}, \mathcal{D}) = \pm 1,$$

$\boldsymbol{G}$ has at least one zero in $\mathcal{D}$. Even more ([1, p. 52]), if (15) is applied to $\boldsymbol{H}$, the Brouwer's degree is invariant for all $\lambda \in [0, 1]$, provided

$$\boldsymbol{H}(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda) \neq \boldsymbol{0}, \quad \boldsymbol{t} \in \partial\mathcal{D},\ \lambda \in [0, 1]. \tag{16}$$

It is also important to note that if $J(\boldsymbol{G})$ in (15) is globally nonsingular, then Brouwer's degree gives the exact number of zeros of $\boldsymbol{G}$ in $\mathcal{D}$. In our case the Jacobian

$$J(\boldsymbol{H})(\boldsymbol{t}) := \frac{\partial \boldsymbol{H}(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda)}{\partial \boldsymbol{t}} = \frac{\partial \boldsymbol{F}(\boldsymbol{t}, \boldsymbol{\alpha}^*)}{\partial \boldsymbol{t}}$$

really is globally nonsingular on $\mathcal{D}$, since its determinant at any point $\boldsymbol{t} \in \mathcal{D}$ simplifies to

$$\frac{6\,(t_0 - t_4)^3\;(t_4 - t_1)\;(t_4 - t_2)\;(t_4 - t_3)}{(t_1 - t_0)\;(t_2 - t_0)\;(t_3 - t_0)\;(t_2 - t_1)^2\,(t_3 - t_1)^2\,(t_3 - t_2)^2} < 0.$$

Since for $\lambda = 0$ the homotopy (14) becomes our particular system (11) for which a unique solution has been established, and

$$\text{degree}(\boldsymbol{H}(\bullet, \boldsymbol{\alpha}^*; \bullet), \mathcal{D}) = -1,$$

the Brouwer's degree of $\boldsymbol{H}$ will be -1 for all $\lambda \in [0,1]$, if (16) holds. Unfortunately, $\boldsymbol{H}$ is not differentiable on $\partial\mathcal{D}$. Even more, it is not continuous and unbounded on some points of the boundary. Thus the following lemma is needed.

Lemma 1.1 *There is a compact set $\widetilde{\mathcal{D}} \subset \mathcal{D}$ which contains particular solution (13) and $\boldsymbol{H}(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda) \neq \boldsymbol{0}$ for $\boldsymbol{t} \in \partial\widetilde{\mathcal{D}}$, $\lambda \in [0,1]$ and $\boldsymbol{\alpha}$ with positive components.*

Proof: Let us prove first that $\boldsymbol{H}(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda)$ can not have any solution arbitrary close to $\partial\mathcal{D}$. Note that

$$t_0 = t_1,\ t_3 < t_4, \quad \text{or} \quad t_0 < t_1,\ t_3 = t_4$$

implies $\boldsymbol{H}_2(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda) \neq 0$, thus $\boldsymbol{H}_2$ is either unbounded or

$$\boldsymbol{H}_2(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda) = (1-\lambda)\boldsymbol{\alpha}_2^* + \lambda\boldsymbol{\alpha}_2 > 0,$$

since the components of $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ are positive. Thus only the relations

$$t_0 = t_1 \leq t_2 \leq t_3 = t_4 \quad \text{or} \quad t_0 < t_1 \leq t_2 \leq t_3 < t_4$$

are left to examine. Since $t_0 = 0 < 1 = t_4$, there are only two possibilities in the first case, $t_0 = t_1 < t_2$ and $t_2 < t_3 = t_4$. This implies obvious contradictions

$$\boldsymbol{H}_1(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda) = (1-\lambda)\boldsymbol{\alpha}_1^* + \lambda\boldsymbol{\alpha}_1 > 0,$$

and

$$\boldsymbol{H}_3(\boldsymbol{t}, \boldsymbol{\alpha}; \lambda) = (1-\lambda)\boldsymbol{\alpha}_3^* + \lambda\boldsymbol{\alpha}_3 > 0.$$

In the second case one has $t_0 < t_1 = t_2 < t_3 < t_4$, $t_0 < t_1 < t_2 = t_3 < t_4$ or $t_0 < t_1 = t_2 = t_3 < t_4$. But now $\boldsymbol{H}_1$ or $\boldsymbol{H}_3$ is unbounded. So all the zeros of $\boldsymbol{H}$ are strictly in $\mathcal{D}$. But $\mathcal{D}$ is an open set, thus there is a compact set $\widetilde{\mathcal{D}} \subset \mathcal{D}$ with a smooth boundary which contains all the zeros of $\boldsymbol{H}$ in its interior.

The proof of the last lemma also completes the proof of the Theorem 1.

4. Numerical example

The results from the previous sections will be illustrated by a numerical example here. Let us suppose that the interpolating points are taken from the helix

$$\boldsymbol{f}(\eta) = [\cos 3\eta, \sin 3\eta, 3\,\eta]^T.$$

Let $\boldsymbol{T}_j = f(\eta_j)$, where $\eta_j = j/4$, $j = 0, 1, 2, 3, 4$. It is a matter of straightforward computation to verify that α_i, $i = 1, 2, 3$, defined by (10), are positive and the conditions of the Theorem 1 are met. The solution of the nonlinear system (9) can be obtained by applying, e.g., Newton's method or any of the continuation methods. This gives the solution

$$\boldsymbol{t} = [0.2313, 0.5000, 0.7687]^T.$$

Now one can use any classical interpolation scheme on arbitrary four interpolation points $\boldsymbol{T}_j$, which gives the interpolating polynomial curve

$$\boldsymbol{B}(t) = \begin{bmatrix} 3.041\,t^3 - 4.823\,t^2 - 0.207\,t + 1 \\ -0.216\,t^3 - 3.384\,t^2 + 3.741\,t \\ 1.172\,t^3 - 1.759\,t^2 + 3.586\,t \end{bmatrix}.$$

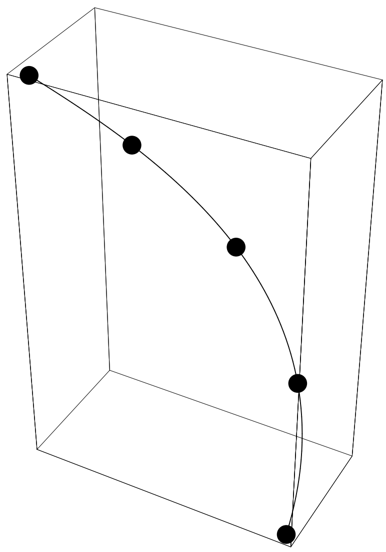

Figure 1. The interpolated data points and parametric polynomial curve.

References

[1] Berger, M. S., *Nonlinearity and Functional Analysis*, Lectures on Nonlinear Problems in Mathematical Analysis, Academic Press, 1977.

[2] de Boor, C., K. Höllig, and M. Sabin, High accuracy geometric Hermite interpolation, Comput. Aided Geom. Design, 4 (1987), pp. 269–278.

[3] Feng, Y. Y., and J. Kozak, On spline interpolation of space data, in Mathematical Methods for Curves and Surfaces II, M. Dæhlen, T. Lyche, and L. L. Schumaker (eds.), Vanderbilt University Press, Nashville, 1998, pp. 167–174.

[4] Höllig, K., and J. Koch, Geometric Hermite interpolation with maximal order and smoothness, Comput. Aided Geom. Design, 13 (1996), pp. 681–695.

[5] Kozak, J., and E. Žagar, On curve interpolation in $\mathbb{R}^d$, in Curve and Surface Fitting - Saint Malo 1999, A. Cohen, C. Rabut, and L. L. Schumaker (eds.), Vanderbilt University Press, Nashville, 2000, pp. 263–273.

[6] Kozak, J., and E. Žagar, On geometric interpolation by polynomial curves, accepted for publication in SIAM Journal on Numerical Analysis.

[7] Mørken, K., Parametric interpolation by quadratic polynomials in the plane, in Mathematical Methods for Curves and Surfaces, M. Dæhlen, T. Lyche, and L. L. Schumaker (eds.), Vanderbilt University Press, Nashville, 1995, pp. 385–402.

[8] Mørken, K., and K. Scherer, A general framework for high-accuracy parametric interpolation, Math. Comp., 66 (1997), pp. 237–260.