Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Cubic Splines with Least-Squares and Conditioned Curvature and Applications

Percy Antonio Ticona Centeno[1], Jorge Lizardo Díaz Calle[2*]

[1] Department of Mathematics, San Augustin National University, Arequipa, Peru
[2] Department of Basic Sciences, Faculty of Animal Science and Food Engineering, University of São Paulo, Brazil
E-mail: jorge.calle@usp.br

Received: 21 October 2024; **Revised:** 27 November 2024; **Accepted:** 11 December 2024

Abstract: In this article is detailed reviewed cubic splines with least-squares method and it is proposed a strategy for the interpolating spline obtained to have a conditioned curvature. The control of this curvature is done from a discrete point of view, resulting in an optimization problem. In the last part, some illustrative experiments and with practical applications have been exposed.

***Keywords*:** least squares, cubic splines, discrete curvature, optimization

MSC: 65L05, 93E24, 34K28

Nomenclature

$M_{n\times m}$	represents the set of arrays of real numbers of $n\times m$ order.
$k\in\mathbb{N}$ and $\ell\in\mathbb{N}$	such that $\ell<(k+1)$ denotes two natural numbers.
I_n	represents the square identity matrix of $n\times n$ order.
$E(x)$	Expected value of x.
$var(x)$	Variance of x.
$S_\Delta(x)$	A cubic polynomial specified over a certain interval, constituting a component of the cubic spline.
$\kappa_\alpha(t)$	The curvature of the vector function α at the parameter t.
$\kappa_f(t)$	Curvature of the real function f at the point t.
κ_i	Discrete curvature at the coordinates $(x_i;y_i)$.
ω	Upper limit for curvature.
$R=\dfrac{1}{\omega}$	Radius of the osculating circle.

1. Introduction

It's undeniable that the least squares strategy has a wide variety of applications in various areas. Besides that, this traditional tool appears in recent impactful areas like data science [1]. Not only to fit using a system of linear equations,

DOI: https://doi.org/10.37256/cm.6220255953

new algorithms also allow solving systems of linear equations in distributed setting [2, 3]. In numerical simulations, it is mentioned that mixed variational formulation for the reaction-diffusion problem based on a saddle point least square [4]. Based on results in [5] it is intended to use the least square technique to apply *hp*-adaptivity in finite element algorithms.

On the other hand, interpolating cubic splines appear in different works linked to design engineering [6], water resources [7], numerical simulation of mathematical models [8–10].

Cubic splines allow a set of discrete data to be interpolated into a polynomial function piece by piece, not simply continuous, but with differentiable properties. However, when the number of data is big, this representation can not satisfy the expectations. For this reason, it's important that those data are represented by curves with less demanding properties, and this is achieved by combining splines with least squares method.

Even when fitting a set of data through cubic splines with least squares, the curvature of this function can not satisfy the needs of the problem. In practical applications it's important to condition that the obtained spline has a superiorly limited curve in each point.

This work takes as a starting point the postulate by Poirier [11], in which is combined and analyzed the cubic splines and the least squares strategy in its discrete and linear version. In addition to this, we seek to fit a set of data with a cubic spline which curvature in each point is superiorly limited.

In the second section, a summary of the most important aspects linked to the least squares method, linear and discrete cases, will be performed. Part of this information has been obtained from [12]. In the third section some important aspects related to interpolating cubic splines will be approached. The fourth section is intended to treat cubic splines with least squares. The extracted information for the third and fourth sections is aligned with the provided approach by Poirier [11]. In the fifth section it's extended the former for the pondered case. In the sixth section a notion of discrete curvature is established. The seventh and eighth sections are reserved to pose and solve a problem of curve fitting through cubic spline with least square and conditioned curvature. Finally, in the last section, a real application problem arises and a simplified version of it is solved.

2. Least squares

The least squares problem is considered for a linear and discrete case. Let be $k+1$ real numbers $x_i \in \mathbb{R}$, $i = 0, 1, \cdots, k$, with $x_0 < x_1 < \cdots < x_k$ and be a real function $f : \mathbb{R} \mapsto \mathbb{R}$ in which $k+1$ images are known and denoted by $y_i = f(x_i)$.

Assuming that ℓ predefined real functions $g_j : \mathbb{R} \mapsto \mathbb{R}$ exist, $j = 1, 2, \cdots, \ell$, the real function $g : \mathbb{R} \mapsto \mathbb{R}$ is constructed, with a linear combination of them:

$$g(x) = \sum_{j=1}^{\ell} \alpha_j g_j(x) \tag{1}$$

in which $\alpha_j \in \mathbb{R}$.

The coefficients α_j, are calculated seeking that g fits to f in the $(k+1)$ nodes x_i. That is to say, seeking to satisfy the overdetermined system of $(k+1)$ linear equations:

$$g(x_i) = \sum_{j=1}^{\ell} \alpha_j g_j(x_i) = y_i \tag{2}$$

The equation (2) is, in matrix form,

$$W\alpha = y \tag{3}$$

where

$$W = [w_{ij} = g_j(x_i)]_{\substack{i=0,\ 1,\ \cdots,\ k \\ j=1,\ 2,\ \cdots,\ \ell}} \quad \alpha = [\alpha_j]_{j=1,\ 2,\ \cdots,\ \ell} \quad \text{and} \quad y = [y_i]_{i=0,\ 1,\ \cdots,\ k}$$

As is known, the overdetermined system (3) frequently doesn't have solution.

2.1 *A deterministic approach*

Solving (3) in the sense of least squares means finding the best α, denoted by $\hat{\alpha}$, so that $W\alpha$ is as near as possible to y. Formally, this proximity is understood by solving the least squares problem

$$\text{Minimize } \frac{1}{2}\|W\alpha - y\|^2 \quad \text{Subject to: } \alpha \in M_{\ell\times 1} \tag{4}$$

in which $\|\cdot\|$ is the Euclidean norm, which solution can be found by solving, at the same time, the system of normal equations

$$W^T W\alpha = W^T y \tag{5}$$

If y is known, this system always has solution [13]. If the range of W is complete, the solution is unique and it's given by

$$\hat{\alpha} = \left(W^T W\right)^{-1} W^T y \tag{6}$$

2.2 *A statistic approach*

Regarding (3), consider now the linear regression model

$$y = W\alpha + \varepsilon \tag{7}$$

in which y and ε are stochastic variables, and $\varepsilon \in M_{(k+1)\times 1}$ is denominated disturbance vector.

In these conditions, the vector $\hat{\alpha}$ obtained in (6), represents an estimator (in the sense of least squares) for α.

It's understood by *bias* to be a systematic error that leads to quantitatively incorrect findings. The biases represent the difference between what is being valued and what is believed to be valued [14]. The analysis of the model (7) depends on the hypothesis that are made about the random variable ε.

To be able to successfully apply the model (7) in statistics, there are two conditions that are fundamental to ensure good properties of the estimator $\hat{\alpha}$ obtained in (7), those are:

1. Normal distribution of the residuals. The residuals must be normally distributed with zero mean, $E(\varepsilon) = \mathbf{0}$. That is to say, the disturbance must be unbiased.

2. Constant variability of the residuals. The variance of the residuals must be constant within all the range of observations. It means that the residuals are distributed in a random way maintaining the same dispersion σ and with no specific pattern, $var(\varepsilon) = \sigma^2 I_{k+1}$. In other words, the disturbance should be homoscedastic.

With those requirements is expected that the estimator $\hat{\alpha}$ is, at least, unbiased. Calculating

$$E(\hat{\alpha}) = E\left(\left(W^T W\right)^{-1} W^T y\right) = E\left(\left(W^T W\right)^{-1} W^T (W\alpha + \varepsilon)\right)$$

$$= E\left(\left(W^T W\right)^{-1} W^T W\alpha\right) + E\left(\left(W^T W\right)^{-1} W^T \varepsilon\right)$$

$$= E(\alpha) + \left(W^T W\right)^{-1} W^T E(\varepsilon) = \alpha + \left(W^T W\right)^{-1} W^T \mathbf{0} = \alpha$$

$\hat{\alpha}$ is unbiased, because $E(\hat{\alpha}) = \alpha$, that is to say, the mean of the estimator coincides with the true populational parameter. In the next section will be demonstrated that, of all the unbiased estimators, $\hat{\alpha}$ is the best.

On the other hand

$$var(\hat{\alpha}) = E\left((\hat{\alpha} - E(\hat{\alpha}))(\hat{\alpha} - E(\hat{\alpha}))^T\right) = E\left((\hat{\alpha} - \alpha)(\hat{\alpha} - \alpha)^T\right)$$

$$= E\left(\left((W^T W)^{-1} W^T (W\alpha + \varepsilon) - \alpha\right)\left((W^T W)^{-1} W^T (W\alpha + \varepsilon) - \alpha\right)^T\right)$$

$$= E\left(\left((W^T W)^{-1} W^T \varepsilon\right)\left((W^T W)^{-1} W^T \varepsilon\right)^T\right) = \left((W^T W)^{-1} W^T E\left(\varepsilon\varepsilon^T\right) W (W^T W)^{-1}\right)$$

$$= \left((W^T W)^{-1} W^T E\left((\varepsilon - \mathbf{0})(\varepsilon - \mathbf{0})^T\right) W (W^T W)^{-1}\right)$$

$$= \left((W^T W)^{-1} W^T var(\varepsilon) W (W^T W)^{-1}\right) = \sigma^2 (W^T W)^{-1}$$

That is to say, $\hat{\alpha}$ is not necessarily homoscedastic.

2.2.1 *The Gauss-Markov theorem*

The Gauss-Markov theorem states that, the linear least squares estimator obtained in (6), is the best linear unbiased estimator. For more information, see [12]. To prove it, it's assumed that there is another unbiased estimator $\bar{\alpha}$. So,

$$\bar{\alpha} = \bar{W} y \tag{8}$$

for some matrix $\bar{W} \in M_{\ell \times (k+1)}$. Then,

$$\bar{\alpha} = \bar{W} y = \bar{W}(W\alpha + \varepsilon) = \bar{W} W\alpha + \bar{W}\varepsilon. \tag{9}$$

As $\bar{\alpha}$ is unbiased,

$$\alpha = E(\bar{\alpha}) = E(\bar{W} W\alpha + \bar{W}\varepsilon) = E(\bar{W} W\alpha) + E(\bar{W}\varepsilon) = \bar{W} W\alpha + \bar{W} E(\varepsilon) = \bar{W} W\alpha + \bar{W}(\mathbf{0}) = \bar{W} W\alpha$$

Then, $\bar{W}W = I_\ell$, so $\bar{\alpha} = \alpha + \bar{W}\varepsilon$, for α not zero.
Now, calculating the variance-covariance matrix of $\bar{\alpha}$, one has

$$var(\bar{\alpha}) = E\left((\bar{\alpha}-\alpha)(\bar{\alpha}-\alpha)^T\right) = E\left((\alpha+\bar{W}\varepsilon-\alpha)(\alpha+\bar{W}\varepsilon-\alpha)^T\right)$$

$$= E\left((\bar{W}\varepsilon)(\bar{W}\varepsilon)^T\right) = E\left(\bar{W}\varepsilon\varepsilon^T\bar{W}^T\right) = \bar{W}E\left(\varepsilon\varepsilon^T\right)\bar{W}^T = \bar{W}\sigma^2 I_{k+1}\bar{W}^T = \sigma^2\bar{W}\bar{W}^T$$

Also

$$var(\bar{\alpha}) - var(\hat{\alpha}) = \sigma^2\bar{W}\bar{W}^T - \sigma^2\left(W^TW\right)^{-1} = \sigma^2\left(\bar{W}\bar{W}^T - \bar{W}W\left(W^TW\right)^{-1}W^T\bar{W}^T\right)$$

$$= \sigma^2\left(\bar{W}\left(I_{k+1} - W\left(W^TW\right)^{-1}W^T\right)\bar{W}^T\right)$$

As well as

$$\left(I_{k+1} - W\left(W^TW\right)^{-1}W^T\right)\left(I_{k+1} - W\left(W^TW\right)^{-1}W^T\right) = I_{k+1} - W\left(W^TW\right)^{-1}W^T \tag{10}$$

it's observed that the matrix

$$I_{k+1} - W\left(W^TW\right)^{-1}W^T \tag{11}$$

is idempotent, and it's not hard to verify that is symmetrical.
Then,

$$var(\bar{\alpha}) - var(\hat{\alpha}) = \sigma^2\bar{W}\left(I_{k+1} - W\left(W^TW\right)^{-1}W^T\right)\left(\bar{W}\left(I_{k+1} - W\left(W^TW\right)^{-1}W^T\right)\right)^T. \tag{12}$$

As the matrix

$$\bar{W}\left(I_{k+1} - W\left(W^TW\right)^{-1}W^T\right)\left(\bar{W}\left(I_{k+1} - W\left(W^TW\right)^{-1}W^T\right)\right)^T \tag{13}$$

is symmetrical and positive semi-defined, then the elements of its diagonal are not negative. Therefore, the variances of the unbiased estimator $\bar{\alpha}$ are greater than or equal the variances of the unbiased estimator $\hat{\alpha}$ provided by least squares.

2.2.2 *Estimator for σ^2*

Although the matrix W is known (the regressors), frequently the value of σ^2 it is not. To estimate σ^2 one proceeds to use the residuals

$$e = y - W\hat{\alpha} = y - W\left(W^T W\right)^{-1} W^T y = \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) y \tag{14}$$

As $y = W\alpha + \varepsilon$, we have

$$e = \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right)(W\alpha + \varepsilon)$$

$$= \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) W\alpha + \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon$$

$$= 0\alpha + \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon = \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon$$

because

$$\left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) W = \mathbf{0}$$

Then,

$$e^T e = \varepsilon^T \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right)^T \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon$$

$$= \varepsilon^T \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon$$

because of the symmetry and idempotence of the matrix in brackets. Therefore,

$$E(e^T e) = E\left(\varepsilon^T \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon\right) = E\left(tr\left[\varepsilon^T \left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon\right]\right)$$

$$= E\left(tr\left[\left(I_{k+1} - W\left(W^T W\right)^{-1} W^T\right) \varepsilon\varepsilon^T\right]\right)$$

$$= tr\left[I_{k+1} - W\left(W^T W\right)^{-1} W^T\right] E\left(\varepsilon\varepsilon^T\right) = tr\left[I_{k+1} - W\left(W^T W\right)^{-1} W^T\right] \sigma^2 I_1$$

$$= \sigma^2\left(tr\left(I_{k+1}\right) - tr\left(W\left(W^T W\right)^{-1} W^T\right)\right) = \sigma^2\left(k + 1 - tr\left(\left(W^T W\right)^{-1} W^T W\right)\right)$$

$$= \sigma^2\left(k + 1 - tr\left(I_\ell\right)\right) = \sigma^2\left(k - \ell + 1\right)$$

Now, making

$$s^2 = \frac{e^T e}{k - \ell + 1} \tag{15}$$

it's obtained that

$$E(s^2) = E\left(\frac{e^T e}{k - \ell + 1}\right) = \frac{1}{k - \ell + 1} E\left(e^T e\right) = \frac{\sigma^2 (k - \ell + 1)}{k - \ell + 1} = \sigma^2 \tag{16}$$

It is concluded that (15) is a biased estimator of the populational variance, given that $E(s^2) = \sigma^2$, and besides, is an estimator of the populational variance of the disturbances.

Representing, the entrances of the reverse matrix, by

$$d_{j,\ \bar{j}} = \left[\left(W^T W\right)^{-1}\right]_{j,\ \bar{j}},\ j,\ \bar{j} = 1,\ 2,\ ...,\ \ell \tag{17}$$

one has that $var(\hat{\alpha}_j) = s^2 d_{j,\ j}$, while $cov(\hat{\alpha}_j,\ \hat{\alpha}_{\bar{j}}) = s^2 d_{j,\ \bar{j}}$.

2.2.3 *Weighted least squares*

Consider the i-th equation of the system in (7), $i = 0,\ 1,\ \cdots,\ k$,

$$y_i = \sum_{j=1}^{\ell} \alpha_j w_{i,\ j} + \varepsilon_i \tag{18}$$

and the $(k+1)$ positive real numbers δ_i (i-th weight), multiplying the i-th equation

$$\delta_i y_i = \delta_i \sum_{j=1}^{\ell} \alpha_j w_{i,\ j} + \delta_i \varepsilon_i \tag{19}$$

Using the following notation $\bar{y}_i = \delta_i y_i$ y $\bar{\varepsilon}_i = \delta_i \varepsilon_i$, it's written

$$\bar{y}_i = \sum_{j=1}^{\ell} \delta_i \alpha_j w_{i,\ j} + \bar{\varepsilon}_i \tag{20}$$

If ε is unbiased, then $\bar{\varepsilon}_i$ is likewise unbiased, since

$$E(\bar{\varepsilon}_i) = E\left(\delta_i \varepsilon_i\right) = \delta_i E\left(\varepsilon_i\right) = \mathbf{0} \tag{21}$$

On the other hand,

$$var(\bar{\varepsilon}_i) = var(\delta_i\varepsilon_i) = \delta_i^2 var(\varepsilon_i) = \delta_i^2\sigma^2 \tag{22}$$

This suggests that homoscedasticity can be lost prior to weighing.

2.2.4 *Generalized least squares*

Consider the i-th equation of the system in (7), $i = 0,\ 1,\ \cdots,\ k$,

$$y_i = \sum_{j=1}^{\ell} \alpha_j w_{i,\ j} + \varepsilon_i \tag{23}$$

If $E(\varepsilon_i) = 0$ and $var(\varepsilon_i) = \sigma^2 u_i$, where u_i is some real not null variable, it's the heteroscedastic case, then the properties for the estimator $\hat{\alpha}$ obtained through least squares can not be favorable. However, if the i-th equation is scaled, in the following manner

$$\frac{y_i}{\sqrt{u_i}} = \frac{\alpha_1}{\sqrt{u_i}} g_1(x_i) + \frac{\alpha_2}{\sqrt{u_i}} g_2(x_i) + \ldots + \frac{\alpha_\ell}{\sqrt{u_i}} g_\ell(x_i) + \frac{\varepsilon_i}{\sqrt{u_i}},\ u_i \neq 0 \tag{24}$$

the new disturbance will be $\varepsilon_i^* = \dfrac{\varepsilon_i}{\sqrt{u_i}}$.

Notice that

$$E(\varepsilon_i^*) = E\left(\frac{\varepsilon_i}{\sqrt{u_i}}\right) = \frac{1}{\sqrt{u_i}} E(\varepsilon_i) = 0,\ u_i \neq 0$$

and

$$var(\varepsilon_i^*) = var\left(\frac{\varepsilon_i}{\sqrt{u_i}}\right) = \left(\frac{1}{\sqrt{u_i}}\right)^2 var(\varepsilon_i) = \frac{1}{u_i} var(\varepsilon_i) = \frac{1}{u_i}(\sigma^2 u_i) = \sigma^2,\ u_i \neq 0$$

That is to say, $var(\varepsilon_i^*) = cte$.

3. Interpolant cubic spline

Let be $k+1$ known real numbers, $x_i \in \mathbb{R}$ with $x_i < x_{i+1}$, $i = 0,\ 1,\ \cdots,\ k$. Let be a real function $f : \mathbb{R} \mapsto \mathbb{R}$, which $k+1$ images are known and denoted by $y_i = f(x_i)$.

According with the approach proposed by Poirier [11], the interpolant cubic spline $S_\Delta : \mathbb{R} \mapsto \mathbb{R}$ for a given $\{x_i\}$ nodes, regarding to each subinterval $[x_{j-1},\ x_j]$, $j = 1,\ 2,\ \ldots,\ k$, is given by

$$S_\Delta(x) = \frac{x_j - x}{6h_j}\left((x_j - x)^2 - h_j^2\right) M_{j-1} + \frac{x - x_{j-1}}{6h_j}\left((x - x_{j-1})^2 - h_j^2\right) M_j + \left(\frac{x_j - x}{h_j}\right) y_{j-1} + \left(\frac{x - x_{j-1}}{h_j}\right) y_j \tag{25}$$

where $h_j = x_j - x_{j-1}$.

The unknowns $\{M_i,\ i = 0,\ 1,\ \cdots,\ k\}$, in (25), results from the solution of the system

$$(1-\lambda_j)M_{j-1} + 2M_j + \lambda_j M_{j+1} = \frac{6y_{j-1}}{h_j(h_j + h_{j+1})} - \frac{6y_j}{h_j h_{j+1}} + \frac{6y_{j+1}}{h_{j+1}(h_j + h_{j+1})} \tag{26}$$

where

$$\lambda_j = \frac{h_{j+1}}{h_j + h_{j+1}}$$

for $j = 1,\ \ldots,\ k-1$.

As can be seen, the system (26) is of the order $(k-1)\times(k+1)$.

The final conditions are given by $M_0 = \pi_0 M_1$ ($|\pi_0| < 2$) and $M_k = \pi_k M_{k-1}$ ($|\pi_k| < 2$), here, it is worked with the natural spline, which means adopting $\pi_0 = 0$ and $\pi_k = 0$. For other alternative elections in the final conditions, see [11].

Let be

$$M = M_{(k+1)\times 1} = [M_i]_{i=0,\ 1,\ \cdots,\ k}, \qquad Y = Y_{(k+1)\times 1} = [y_i]_{i=0,\ 1,\ \cdots,\ k}$$

and the tridiagonal matrix $A_{(k+1)\times(k+1)}$,

$$A = \begin{bmatrix} a_{1,\,1} & a_{1,\,2} & 0 & \cdots & 0 & 0 & 0 \\ a_{2,\,1} & a_{2,\,2} & a_{2,\,3} & \cdots & 0 & 0 & 0 \\ 0 & a_{3,\,2} & a_{3,\,3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{k-1,\,k-1} & a_{k-1,\,k} & 0 \\ 0 & 0 & 0 & \cdots & a_{k,\,k-1} & a_{k,\,k} & a_{k,\,k+1} \\ 0 & 0 & 0 & \cdots & 0 & a_{k+1,\,k} & a_{k+1,\,k+1} \end{bmatrix}$$

where

$$\begin{cases} a_{j,\,j} = 2, & j = 1,\ 2,\ \ldots,\ k+1 \\ a_{1,\,2} = -2\pi_0 & \\ a_{j,\,j+1} = \lambda_{j-1}, & j = 2,\ \ldots,\ k \\ a_{j+1,\,j} = 1-\lambda_j, & j = 1,\ \ldots,\ k-1 \\ a_{k+1,\,k} = -2\pi_k & \end{cases}$$

So too, the tridiagonal matrix $\theta_{(k+1)\times(k+1)}$,

$$\theta = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \theta_{2,\,1} & \theta_{2,\,2} & \theta_{2,\,3} & \cdots & 0 & 0 & 0 \\ 0 & \theta_{3,\,2} & \theta_{3,\,3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \theta_{k-1,\,k-1} & \theta_{k-1,\,k} & 0 \\ 0 & 0 & 0 & \cdots & \theta_{k,\,k-1} & \theta_{k,\,k} & \theta_{k,\,k+1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

where

$$\begin{cases} \theta_{j,\,j} = -\dfrac{6}{h_{j-1}h_j} & j = 2,\ 3,\ ...,\ k \\ \theta_{j-1,\,j} = \dfrac{6}{h_{j-1}(h_{j-2}+h_{j-1})} & j = 3,\ 4,\ ...,\ k+1 \\ \theta_{j,\,j-1} = \dfrac{6}{h_{j-1}(h_{j-1}+h_j)} & j = 2,\ 3,\ ...,\ k \end{cases}$$

Then, M can be found by solving the linear equation system

$$AM = \theta Y \tag{27}$$

When Y is known, and being A diagonally dominant, the solution of (27) can be found by making

$$M = A^{-1}\theta Y \tag{28}$$

Now, using an arbitrary n-vector

$$\xi = [\xi_j]_{j=1,\,2,\,\cdots,\,n}$$

it's seen that for this corresponds to a values n-vector

$$S_\Delta(\xi) = [S_\Delta(\xi_j)]_{j=1,\,2,\,\cdots,\,n}.$$

of spline interpolants as a linear function of the vector Y.

To obtain a matrix formulation of $S_\Delta(\xi)$ in terms of Y, it's necessary defining a few coefficients matrices.

Considering (25), let be

$$P = P_{n\times(k+1)} = [p_{jm}]_{\substack{j=1,\,2,\,\cdots,\,n \\ m=0,\,1,\,\cdots,\,k}} \quad \text{and} \quad Q = Q_{n\times(k+1)} = [q_{jm}]_{\substack{j=1,\,2,\,\cdots,\,n \\ m=0,\,1,\,\cdots,\,k}}$$

two matrices, such that, for $x_{i-1} \leq \xi_j \leq x_i$, where $i = 1,\ 2,\ \ldots,\ k$ and $j = 1,\ 2,\ \ldots,\ n$,

$$p_{jm} = \begin{cases} \dfrac{x_i - \xi_j}{6h_i}\left[(x_i - \xi_j)^2 - h_i^2\right] & \text{if } m = i-1 \\ \dfrac{\xi_j - x_{i-1}}{6h_i}\left[(\xi_j - x_{i-1})^2 - h_i^2\right] & \text{if } m = i \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

and

$$q_{jm} = \begin{cases} \dfrac{x_i - \xi_j}{h_i} & \text{if } m = i-1 \\ \dfrac{\xi_j - x_{i-1}}{h_i} & \text{if } m = i \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

where $h_i = x_i - x_{i-1}$. Then, using (25), (28-30), satisfies

$$S_\Delta(\xi) = PM + QY = \left(PA^{-1}\theta + Q\right)Y = WY \tag{31}$$

where

$$W = PA^{-1}\theta + Q \tag{32}$$

That is to say, for any n-vector represented by ξ, using W can already obtain directly (without exhibiting explicitly the cubic spline) its corresponding images, $S_\Delta(\xi)$. That way, the pairs $(\xi_j;\ S_\Delta(\xi_j))$ are on the mentioned cubic spline.

Example 1 Consider

$$x = [-1\ \ 1\ \ 3\ \ 5]^T \quad y = [1.0\ \ 1.5\ \ 2.9\ \ 1.8]^T$$

and

$$\xi = [-0.9\quad 0.5\quad 0.1\quad 0.7\quad 1.3\quad 1.9\quad 2.3\quad 2.7\quad 3.1\quad 3.5\quad 3.9\quad 4.2\quad 4.6\quad 4.9]^T$$

Then, making $\eta = Wy$, we have

$$\eta = [1.00\quad 1.24\quad 1.12\quad 1.33\quad 1.72\quad 2.23\quad 2.56\quad 2.80\quad 2.91\quad 2.86\quad 2.68\quad 2.48\quad 2.16\quad 1.89]^T$$

So, the points $(\xi_j;\ \eta_j)$, $j = 1,\ 2,\ \ldots,\ 14$, are on the cubic spline graphic that interpolates the $(-1;\ 1.0)$, $(1;\ 1.5)$, $(3;\ 2.9)$ and $(5;\ 1.8)$ points, as it's seen in the Figure 1.

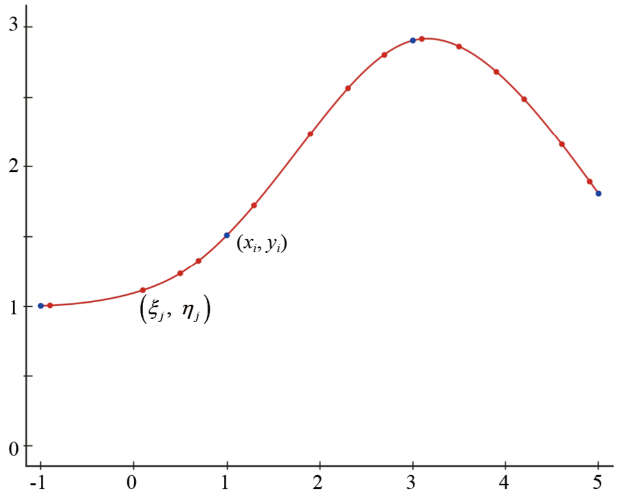

Figure 1. The points (ξ_j, η_j) on the interpolating cubic spline, $j = 1, 2, \ldots, 14$

4. Cubic spline with least squares

Given a partition $\Delta = \{x_i / i = 0, 1, \cdots, k\}$ of the interval $[x_0, x_k]$. Let be $Y = [y_i]^T$ is a $(k+1)$-vector of unknowns corresponding to the nodes of Δ. Be also $\xi = [\xi_j]^T$ and $\eta = [\eta_j]^T$ the n-vectors of known data associated to the independent and dependent variables, respectively, with $n \geq k+1$, where $j = 1, 2, \ldots, n$.

So it's also assumed that the matrix W, given in (32), have range $k+1$.

Lastly, let be $\varepsilon = [\varepsilon_j]^T$ an independent disturbance vector normally distributed, such that

$$E(\varepsilon) = 0 \quad \text{and} \quad E(\varepsilon\varepsilon^T) = \sigma^2 I_n$$

where 0 is a null n-vector. How it's mentioned on the 2.2 subsection, this last is important to obtain significative results.

Thus, the cubic spline regression model will be

$$S_\Delta(\xi) = WY \tag{33}$$

The data $(\xi_j; \eta_j)$, provided in ξ and η, satisfy the cubic spline regression model if, and only if,

$$S_\Delta(\xi) + \varepsilon = WY + \varepsilon = \eta \tag{34}$$

Solving (34) in the sense of least squares, is obtained $\hat{Y}$, the estimate of Y, through:

$$\hat{Y} = \left(W^T W\right)^{-1} W^T \eta \tag{35}$$

That is to say, through this process, given the vectors of data ξ and η, in addition of the predefined nodes x_i, using (35) one can find values for y_i that makes the cubic spline fit on the points $(\xi_j; \eta_j)$ and interpolate the points $(x_i; y_i)$.

Example 2 Consider

$$x = [\ 0 \quad 1 \quad 2 \quad 3 \quad 4\ ]^T$$

$$\xi = [\ 0.1 \quad 0.5 \quad 0.8 \quad 1.2 \quad 1.5 \quad 1.7 \quad 1.9 \quad 2.3 \quad 2.6 \quad 2.9 \quad 3.1 \quad 3.4 \quad 3.7 \quad 3.9\ ]^T$$

$$\eta = [\ 3.1 \quad 2.9 \quad 2.6 \quad 2.5 \quad 2.6 \quad 2.8 \quad 3.1 \quad 3.4 \quad 3.3 \quad 3.4 \quad 3.2 \quad 2.9 \quad 2.7 \quad 2.5\ ]^T \tag{36}$$

According to (35),

$$\hat{Y} = [\ 3.3078 \quad 2.5038 \quad 3.1086 \quad 3.2828 \quad 2.3424\ ]^T$$

So, with x and $\hat{Y}$, using (25), it's already possible to build the cubic spline that interpolates the points $(x_j;\ \hat{Y}_j)$, $j = 0,\ 1,\ 2,\ 3,\ 4$. The continuous line represents the cubic spline interpolating the 5 points (blue) and fitting to the 14 points (black), see the Figure 2.

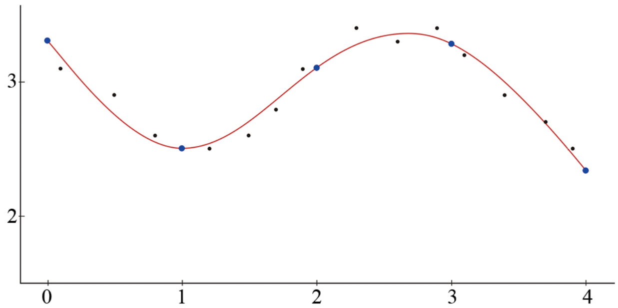

Figure 2. Cubic spline interpolation of the points $(x_j;\ \hat{Y}_j)$, $j = 0,\ 1,\ 2,\ 3,\ 4$

5. Cubic splines with weighted least squares

Let be $d_{j,\ j} > 0$ weights related to the overdetermined system equations seen in (34), where $j = 1,\ 2,\ \cdots,\ n$. If $D \in M_{n\times n}$ is a diagonal matrix which entrances are precisely d_{jj}, then the system to be solved is

$$DWY + D\varepsilon = D\eta \tag{37}$$

The solution, in the sense of least squares is obtained through

$$\tilde{Y} = \left(W^T D^2 W\right)^{-1} W^T D^2 \eta \tag{38}$$

Example 3 Regarding the data in (36), to weight the nodes $x_1 = 0.5$ and $x_2 = 0.8$ we create the diagonal weight matrix $D_{14\times14}$, where $d_{22} = 1,000$, $d_{33} = 1,000$ y $d_{ii} = 1$, $i = 1,\ 4,\ 5,\ \ldots,\ 14$. Therefore, according to (38),

$$\tilde{Y} = [\ 3.645721 \quad 2.5087 \quad 3.1014 \quad 3.2840 \quad 2.3421\ ]^T$$

The Figure 3, illustrates the weighted case (solid line) and the conventional (dashed line).

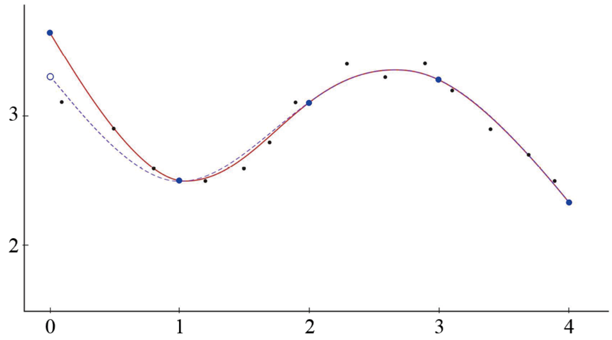

Figure 3. Weighted fitting (solid line) versus conventional fitting (dashed line)

6. Notion of discrete curvature

Be a curve $\alpha : \mathbb{R} \mapsto \mathbb{R}^2$, where $\alpha(t) = (x(t),\ y(t))$ sufficiently smooth. The κ curvature in t of α has various characterizations. One of which for example

$$\kappa_\alpha(t) = \frac{x'(t)y''(t) - x''(t)y'(t)}{\left((x'(t))^2 + (y'(t))^2\right)^{\frac{3}{2}}} \tag{39}$$

Consequently, one can say that, if $f : \mathbb{R} \mapsto \mathbb{R}$, the curvature of f in x it's given by

$$\kappa_f(x) = \frac{f''(x)}{(1 + (f'(t))^2)^{\frac{3}{2}}} \tag{40}$$

However, when discrete data is used, this concept should be addressed in a different way.

Consider a i-set of three points ordered sequentially, on a curve α

$$A = (x_{i-1};\ y_{i-1}),\ B = (x_i;\ y_i) \text{ and } \ C = (x_{i+1};\ y_{i+1}) \tag{41}$$

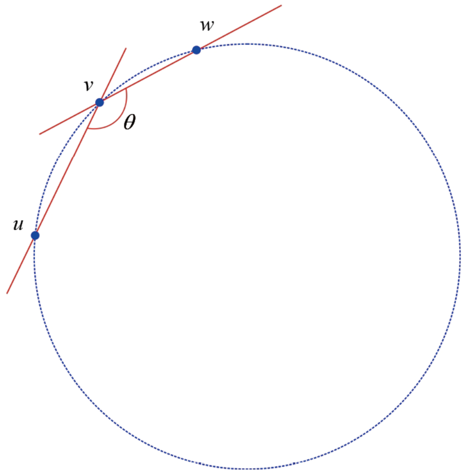

Figure 4. Osculating circle defined by three points

Using the osculating circle (Figure 4), see [15], the discrete curvature regarding the point B, of the curve α that goes through the three points, can be calculated in the following way

$$\kappa_B = \frac{2\sin(\theta)}{\|C-A\|}$$

Notice that θ is the angle between the vectors joining the points

$$\cos\theta = \frac{\langle A-B,\ C-B\rangle}{\|A-B\|\,\|C-B\|}$$

It's known that, if R_B is the radius of the osculating circle, then $\kappa_B = \dfrac{1}{R_B}$.

However, the said circle can be also found in the following way. Consider the equation of the circumference that passes through $(x_{i-1};\ y_{i-1})$, $(x_i;\ y_i)$ and $(x_{i+1};\ y_{i+1})$:

$$x^2+y^2+a_1x+a_2y+a_3=0 \tag{42}$$

Clearly, the center is $\left(-\dfrac{a_1}{2};\ -\dfrac{a_2}{2}\right)$ and the radius is given by $R_B=\sqrt{\dfrac{a_1^2}{4}+\dfrac{a_2^2}{4}-a_3}$.

The values of the coefficients a_1, a_2 and a_3 can be found by solving the system

$$a_1x_{i-1}+a_2y_{i-1}+a_3=-x_{i-1}^2-y_{i-1}^2$$

$$a_1x_i+a_2y_i+a_3=-x_i^2-y_i^2$$

$$a_1x_{i+1}+a_2y_{i+1}+a_3=-x_{i+1}^2-y_{i+1}^2 \tag{43}$$

Notice that (43) has unique solution if the points $(x_{i-1};\ y_{i-1})$, $(x_i;\ y_i)$ and $(x_{i+1};\ y_{i+1})$ are not collinear.

7. Discrete bounded curvature

Consider some i-set of three points on a curve α, as in (41), $(x_{i-1};\ y_{i-1})$, $(x_i;\ y_i)$ and $(x_{i+1};\ y_{i+1})$. The discrete curvature at the vertex $(x_i;\ y_i)$ is assumed to be κ_i.

It's now required to find the points $(x_{i-1};\ \bar{y}_{i-1})$, $(x_i;\ \bar{y}_i)$ and $(x_{i+1};\ \bar{y}_{i+1})$, closest to $(x_{i-1};\ y_{i-1})$, $(x_i;\ y_i)$ and $(x_{i+1};\ y_{i+1})$, respectively, so that the discrete curvature $\bar{\kappa}_i$, regarding the $(x_i;\ \bar{y}_i)$ vertex is superiorly bounded by $\hat{\kappa}$. That is

$$\bar{\kappa}_i \leq \hat{\kappa} \tag{44}$$

Defining $R = 1/\hat{\kappa}$, then if R_i is the radius of the osculating circle relative to the points $(x_{i-1};\ \bar{y}_{i-1})$; $(x_i;\ \bar{y}_i)$ and $(x_{i+1};\ \bar{y}_{i+1})$, so demanding (44) is equivalent to demand

$$R_i \geq R \tag{45}$$

That is to say, as detailed above, to find the osculating circle regarding the points $(x_{i-1};\ \bar{y}_{i-1})$, $(x_i;\ \bar{y}_i)$ and $(x_{i+1};\ \bar{y}_{i+1})$, which $\bar{\kappa}_i$ curvature be superiorly bounded by $\hat{\kappa}$, we should solve a least squares problem with restrictions

$$\begin{aligned}
&\text{Minimize} \quad \frac{\rho}{2}\|y - y_i\|^2 \\
&\text{Subject to:}\ x_{i-1}^2 + \bar{y}_{i-1}^2 + a_1 x_{i-1} + a_2 \bar{y}_{i-1} + a_3 = 0 \\
&x_i^2 + \bar{y}_i^2 + a_1 x_i + a_2 \bar{y}_i + a_3 = 0 \\
&x_{i+1}^2 + \bar{y}_{i+1}^2 + a_1 x_{i+1} + a_2 \bar{y}_{i+1} + a_3 = 0 \\
&a_1^2 + a_2^2 - 4a_3 - 4R^2 \geq 0
\end{aligned} \tag{46}$$

where $\rho > 0$ is a parameter and $\|\cdot\|$ the Euclidean norm.

Controlling the discrete curvature in a curve that passes through more than three points gets complicated. Indeed, be four points

$$(x_{i-1};\ \bar{y}_{i-1}),\ (x_i;\ \bar{y}_i),\ (x_{i+1};\ \bar{y}_{i+1}) \text{ and } (x_{i+2};\ \bar{y}_{i+2})$$

Controlling the discrete curvature in the $(x_i;\ \bar{y}_i)$ vertex using the three first points should be possible, but applying the same principle to control the curvature in $(x_{i+1};\ \bar{y}_{i+1})$, the variables $\bar{y}_i$ and $\bar{y}_{i+1}$ can assume new values, that would modify the discrete curvature already calculated in $(x_i;\ \bar{y}_i)$.

In the next section a way to approach this inconvenient arises, merging this with the cubic splines with least squares strategy.

8. Cubic splines with least squares and conditioned curvature

Let be $x,\ y \in M_{(k+1)\times 1}$, where x is known in a default way, while y is a vector of unknowns to be defined. So also, be the known data $\xi,\ \eta \in M_{n\times 1}$, with $n > k+1$, where at least exists a ξ_j, for some $j = 1,\ 2,\ \ldots,\ n$, belonging to the $[x_{i-1},\ x_i]$, $i = 1,\ 2,\ \ldots,\ k$ interval. For example,

$$x_0 < \xi_1 < \xi_2 < x_1$$

$$x_1 < \xi_3 < \xi_4 < \xi_5 < x_2$$

$$\vdots$$

$$x_{k-1} < \xi_n < x_k$$

Recalling that the cubic spline with least squares that interpolates the pairs $(x_i;\ y_i)$ and fits to the pairs $(\xi_j;\ \eta_j)$, is determined by the system of overdetermined linear equations

$$Wy = \eta \tag{47}$$

where $W \in M_{n\times(k+1)}$ is given in (32). To find the estimated y one should solve (47) in the sense of least squares

$$\bar{y} = (W^T W)^{-1} W^T \eta \tag{48}$$

Once determined the $x,\ \bar{y} \in M_{(k+1)\times 1}$ vectors, one can expose the cubic spline that interpolates the pairs $(x_i;\ \bar{y}_i)$, $i = 0,\ 1,\ \ldots,\ k$, and fits to the pairs $(\xi_j;\ \eta_j)$, $j = 1,\ 2,\ \ldots,\ n$.

However, what happens if it's required that the curvature of the cubic spline is superiorly bounded by a constant ω in $[x_0,\ x_k]$?

Due to the inconveniences that can arise from satisfying this last requirement, it was decided to control the curvature from the discrete point of view.

So, let ω be a positive constant as an upper bound for the curvature of the cubic spline in $[x_0,\ x_k]$. According with the analyzed in the previous section (considering $R = \dfrac{1}{\omega}$), this can be obtained now by solving the following problem

$$\text{Minimize} \quad \frac{\rho_1}{2}\|y-\bar{y}\|^2+\frac{\rho_2}{2}\|Wy-\eta\|^2$$

$$\text{Subject to:} \; x_{i-1}^2+y_{i-1}^2+a_{i,\,1}x_{i-1}+a_{i,\,2}y_{i-1}+a_{i,\,3}=0$$

$$x_i^2+y_i^2+a_{i,\,1}x_i+a_{i,\,2}y_i+a_{i,\,3}=0$$

$$x_{i+1}^2+y_{i+1}^2+a_{i,\,1}x_{i+1}+a_{i,\,2}y_{i+1}+a_{i,\,3}=0$$

$$a_{i,\,1}^2+a_{i,\,2}^2-4a_{i,\,3}-4R^2\geq 0,\; i=1,\,2,\,\ldots,\,k-1 \tag{49}$$

where $\rho_1,\ \rho_2>0$ are weighting parameters, those can be extended in order to favor some components of $Wy-\eta$ or $y-\bar{y}$.

Alternatively, the problem (49) can be seen as

$$\text{Minimize} \quad \frac{\rho_1}{2}\|y-\bar{y}\|^2+\frac{\rho_2}{2}\|Wy-\eta\|^2$$

$$\text{Subject to:} \; x_{i-1}^2+y_{i-1}^2+a_{i,\,1}x_{i-1}+a_{i,\,2}y_{i-1}+a_{i,\,3}=0$$

$$x_i^2+y_i^2+a_{i,\,1}x_i+a_{i,\,2}y_i+a_{i,\,3}=0$$

$$x_{i+1}^2+y_{i+1}^2+a_{i,\,1}x_{i+1}+a_{i,\,2}y_{i+1}+a_{i,\,3}=0$$

$$a_{i,\,1}^2+a_{i,\,2}^2-4a_{i,\,3}-4R^2-d_i=0$$

$$d_i\geq 0,$$

$$i=1,\,2,\,\ldots,\,k-1 \tag{50}$$

Different optimization methods can be employed to resolve problems (49) and (50). It is essential to emphasize the importance of utilizing appropriate initial points. In the performed tests, the curvature data from $(k-1)$ vertices, obtained using the cubic spline interpolant through least squares, was utilized prior to curvature control. That is to say, the values of the coefficients of each osculating circle ($x^2+y^2+ax+by+c=0$) in each vertex $(x_i,\ y_i)$, $i=1,\ \ldots,\ k-1$, as well as the values of $\bar{y}$ obtained through solving (47).

In the following examples, it's used the internal penalty strategy to optimize the problem (50).

Example 4 Consider

$$x = \begin{bmatrix} 0 & 1 & 3 & 5 \end{bmatrix}^T$$

$$\xi = \begin{bmatrix} 0.05 & 0.2 & 0.4 & 0.6 & 0.8 & 0.95 & 1.2 & 1.6 & 1.80 & 1.95 & 2.50 & 2.91 & 3.30 & 3.9 & 4.2 & 4.8 \end{bmatrix}^T$$

$$\eta = \begin{bmatrix} 1.00 & 1.4 & 2.0 & 2.2 & 2.7 & 3.00 & 3.3 & 3.5 & 3.55 & 3.60 & 3.86 & 3.80 & 3.92 & 3.9 & 3.7 & 3.5 \end{bmatrix}^T$$

Then, according to (32), we have

$$\bar{y} = (W^T W)^{-1} W^T \eta = [\ 1.016402 \quad 2.930285 \quad 3.850290 \quad 3.560726\ ]^T$$

Notice that, evaluating $S_\Delta(\xi) = W\bar{y}$, the points $(\xi_j;\ S_\Delta(\xi)_j)$ are on the graphic of the interpolating cubic spline with least squares initially. The discrete curvatures in the vertices $(1;\ 2.930285)$ and $(3;\ 3.850290)$ are 0.2967 and 0.2688, respectively.

If is imposed that the curvatures in those vertices be superiorly bounded by $\omega = 0.25$ (that the radius of the circle be inferiorly bounded by $R = \dfrac{1}{\omega} = 4$), solving the problem (50), for $k = 3$, the new value for y will be

$$y = [\ 1.147869 \quad 2.806620 \quad 3.887879 \quad 3.797090\ ]^T$$

and with that the new interpolating spline is obtained. The discrete curvature in the new vertices $(1;\ 2.806620)$ and $(3;\ 3.887879)$ will be 0.249975 and 0.249941, respectively. See the Figure 5.

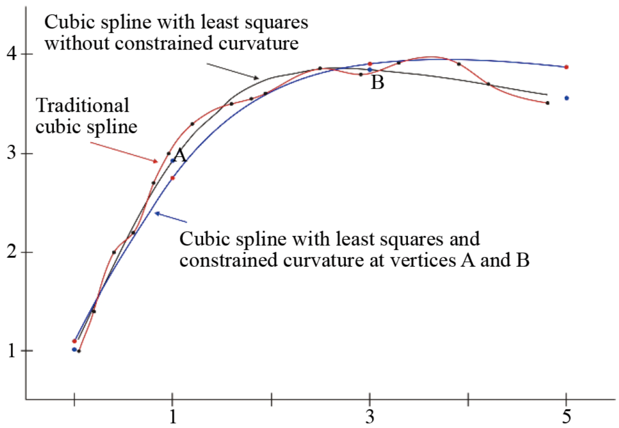

Figure 5. Interpolating cubic spline with constrained curvature by $\omega = 0.25$

Example 5 In the Figure 6, in blue it's shown the interpolating cubic spline (in 6 points) fitting to 24 data. In red, is the case with superiorly bounded curvature by $\omega = 0.37$.

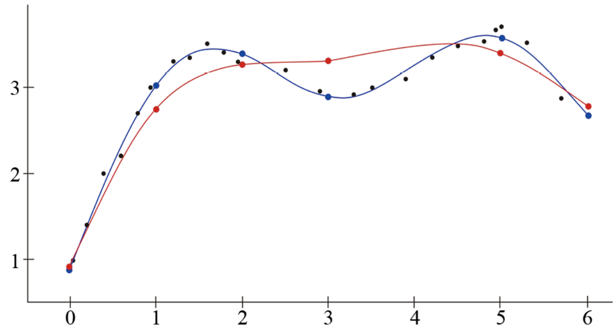

Figure 6. Interpolating cubic splines with constrained curvature by $\omega = 0.37$

Example 6 In the Figure 7, in blue, the interpolating cubic spline (in 8 points) fitting to 26 data. In red, the restricted case to a superiorly bounded curvature by $\omega = 0.3333$, in the sixth vertex. In green, the case for a superiorly bounded curvature by $\omega = 0.5$ in the same sixth vertex.

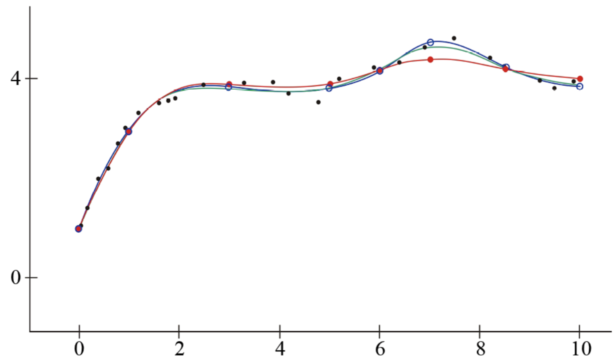

Figure 7. The interpolation cubic splines with constrained curvatures

9. A real application problem 1

In Peru, to quote only a case, the cities Arequipa and Mollendo are connected by a two-way road, which represents a double danger when traveling (specifically when overtaking). This becomes difficult since the road has improperly sized vertical undulations in several places, making it impossible to see the car ahead in the opposite direction because it is on a concavity at that point. With that, rose the necessity of additional information in the maps that appear in the everyday applications, one can obtain information of the vertical undulations of the road in a certain section where the driver is located in real time.

All that remains is to take samples that include the time, latitude, longitude, altitude, among others, that permit us to build a model that be the answer to the question proposed above. Actually, some cell phones come with web capabilities for getting these data.

This problem can be solved using cubic splines with least squares. However, due to the sampling errors, some vertices may show much pronounced curvatures that don't represent the reality, needing the establishment of a superior height for the curvatures in each vertex.

For illustrating purposes, it was chosen to analyze the route that unites two localities within the Arequipa region, a section of the route from Socabaya to Yarabamba.

The following experiment corresponds to a sample of 221 data, relating to the altitudes of a 2 km automotive traffic road. The blue line is associated with the cubic spline that interpolates the 7 blue circles and that fits to the 221 data (black points). The initial curvatures in the vertices 3 and 5 are 0.0100 and 0.01164, respectively. By demanding a maximum curvature of 0.0083, we obtain a cubic spline (in red) as is shown in the Figure 8.

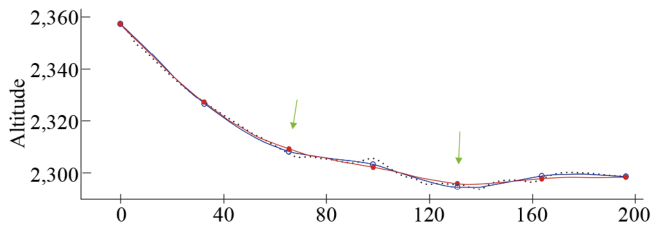

Figure 8. Cubic spline with maximum curvature of $\omega = 0.0083$

10. A real application problem 2

In Table 1 are $n = 59$ data, obtained during 59 seconds, of the latitudes above sea level in a 335.34 m trajectory, in a place of Arequipa.

Table 1. Arequipa: Trajectory of 335.34 meters

Time (s)	Latitude	Longitude	Speed (m/s)	Altitude (m)
1	−16.4495749999	−71.5857769999	0.1320	2,174.8
2	−16.4495730000	−71.5857800000	0.1140	2,175.6
3	−16.4495679999	−71.5857819999	0.0670	2,175.1
4	−16.4495600000	−71.5857830000	0.0340	2,173.6
5	−16.4495569999	−71.5857879999	0.0220	2,172.4
6	−16.4495569999	−71.5857979999	0.1420	2,172.6
7	−16.4495600000	−71.5858080000	0.4000	2,172.8
8	−16.4495770000	−71.5858420000	2.1190	2,173.8
9	−16.4495900000	−71.5858879999	3.4420	2,173.7
⋮	⋮	⋮	⋮	⋮
58	−16.4509800000	−71.5887199999	0.0000	2,186.4
59	−16.4509800000	−71.5887199999	0.0000	2,186.4

Considering the altitude in relation to the meters traveled on the straight line, seeking to fit these data through a cubic spline with least squares, where

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \vdots \\ \xi_{59} \end{bmatrix} = \begin{bmatrix} 0.132 \\ 0.246 \\ 0.313 \\ 0.347 \\ \vdots \\ 335.34 \end{bmatrix}, \; \eta = \begin{bmatrix} \eta_1 \\ \eta_1 \\ \eta_2 \\ \eta_3 \\ \vdots \\ \eta_{59} \end{bmatrix} = \begin{bmatrix} 2,174.8 \\ 2,175.6 \\ 2,175.1 \\ 2,173.6 \\ \vdots \\ 2,186.4 \end{bmatrix}$$

and, considering the nodes

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 28 \\ 56 \\ 84 \\ \vdots \\ 336 \end{bmatrix}$$

one has, applying (35). The searched graphic of the cubic spline is presented in the Figure 9.

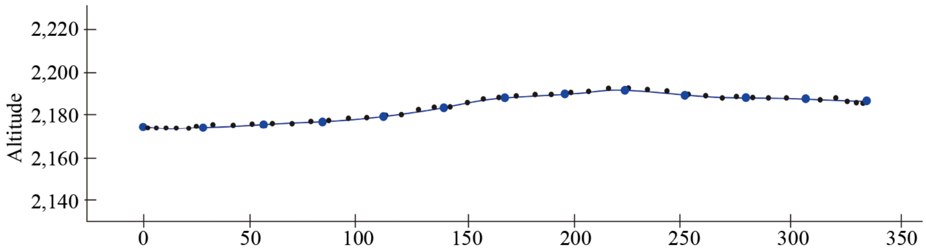

Figure 9. Searched cubic spline in blue line

The curvature in the vertices $(x_1; \bar{y}_1)$, ..., $(x_{11}; \bar{y}_{11})$ are, respectively

0.00118988

0.00030811

0.00156599

0.00183382

0.00045169

0.00389716

0.00085781

0.00569123

0.00131907

0.00122094

$$0.00137851$$

The Figure 10 expands the occurred around the vertex

$$(x_8;\ \bar{y}_8) = (224;\ 2,191.75074698)$$

which discreet curvature is 0.0056912 (radius of the osculating circle 175.708897).

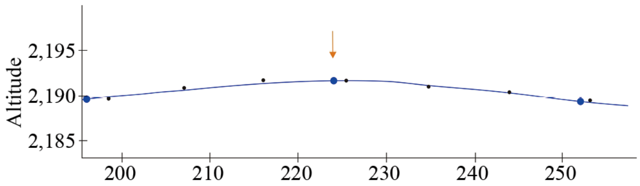

Figure 10. Expanded cubic spline around the vertex $(x_8;\ \bar{y}_8)$

Now, demanding that the curvature in the $(x_8;\ \bar{y}_8)$ be superiorly bounded by 0.0039216 (that the radius of the osculating circle be inferiorly bounded by 255), one has the new searched cubic spline (line in red) (Figure 11), Figure 12 illustrates the distinction between the least square cubic spline and the least square cubic spline with constrained curvature at vertex $(x_8;\ \bar{y}_8)$.

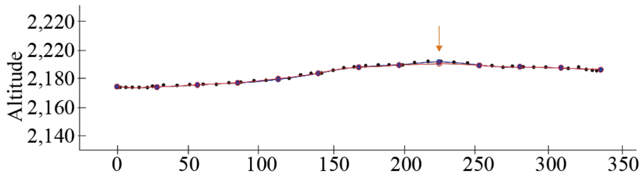

Figure 11. Cubic spline around the vertex $(x_8;\ \bar{y}_8)$ with $\omega = 0.0039216$ (red)

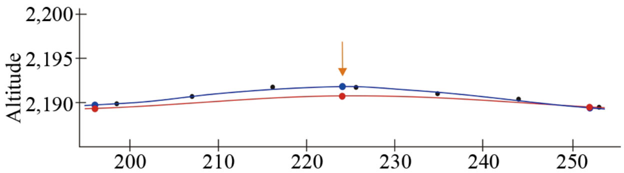

Figure 12. Zooming in the Figure 11 that illustrates the distinction between the least square cubic spline and the least square cubic spline with constrained curvature at vertex $(x_8;\ \bar{y}_8)$

The new curvature in the new vertex $(x_1;\ y_1)$, ..., $(x_{11};\ y_{11})$ are, respectively

0.00118986

0.00030810

0.00156598

0.00142994

0.00045167

0.00389706

0.00018363

0.00325564

0.00024501

0.00055552

0.00137841

It can be seen that there was an essential variation in the second component of the vertex of interest and only small (or null) variations in the second components of the remains of the vertices, as expected. However, the new curvatures, may differ from the previous.

11. Conclusions

The curve fitting through cubic spline with least squares turns to be a very valuable strategy, since combines the sturdiness of a cubic spline with the practicality of least squares. The resulting process is simple and fast to be applied in practice. Can be of great utility when is desired to extrapolate a set of data in different investigation areas.

The curve fitting through least squares and conditioned curvatures is even more valuable, since permits correcting possible fittings where the curvature of the resulting cubic spline doesn't satisfy determined established parameters. The choice of the optimization method and the use of adequate initial points is essential for the solution of the problem (50).

The real application problem presented in the beginning of the section 9, has a social importance. Taking samples of all the trajectory between Arequipa and Mollendo as well as the fitting through cubic splines with least squares with conditioned curvature is viable in practice. The implementation of an application and the extension towards other areas and realities is totally possible.

Acknowledgement

The “Universidad Nacional de San Agustin de Arequipa” is acknowledged by the writers. The second author thankfully acknowledges to the “Universidade de São Paulo”.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Bruce P, Bruce A, Gedeck P. *Estadística Práctica Para Ciencia De Datos Con R Y Python [Practical Statistics for Data Science with R and Python]*. Marcombo Importacion; 2022.

[2] Jahvani M, Guay M. Solving least-squares problems in directed networks: A distributed approach. *Computers and Chemical Engineering*. 2024; 184: 108654.

[3] Wang P, Mou S, Lian J, Ren W. Solving a system of linear equations: From centralized to distributed algorithms. *Annual Reviews in Control*. 2019; 47: 306-322.

[4] Bacuta C, Jacavage J. Saddle point least squares for the reaction-diffusion problem. *Results in Applied Mathematics*. 2020; 8: 100105.

[5] Diaz Calle JL, Pastor JD, Bustincio RW, Ancori R. Study of the enriched mixed finite element method comparing errors and computational cost with classical FEM and mixed scheme on quadrilateral meshes. *Results in Applied Mathematics*. 2021; 10: 100150.

[6] Ramírez CJ. *Acoplamiento aeroelástico de una aeronave completa mediante interpolación con splines [Whole machine aeroelastic coupling based on spline interpolation]*. Universitat Politècnica de València [Universidade Polit écnica de Valencia]; 2023.

[7] Gonzales Zenteno HE. *Predicción del fenómeno El Niño mediante índices oceánicos e influencia de la zona de convergencia intertropical en el norte peruano [Predicting El Nino phenomenon using ocean indices and its impact on the tropical convergence zone in northern Peru]*. Unalm-Escuela de Posgrado [Unalm Graduate School]; 2022.

[8] Negero NT. A uniformly convergent numerical scheme for two parameters singularly perturbed parabolic convection-diffusion problems with a large temporal lag. *Results in Applied Mathematics*. 2022; 16: 100338.

[9] Ayele MA. Fitted cubic spline scheme for two-parameter singularly perturbed time-delay parabolic problems. *Results in Applied Mathematics*. 2023; 18: 100361.

[10] Duressa GF. A robust higher-order fitted mesh numerical method for solving singularly perturbed parabolic reaction-diffusion problems. *Results in Applied Mathematics*. 2023; 20: 100405.

[11] Poirier DJ. Piecewise regression using cubic splines. *Journal of the American Statistical Association*. 1973; 68: 515-524.

[12] Rencher AC, Schaalje GB. *Linear Models in Statistics.* John Wiley and Sons; 2008.

[13] Watkins DS. *Fundamentals of Matrix Computations.* John Wiley and Sons; 2004.

[14] Casal J, Mateu E. Los sesgos y su Control [Prejudice and its Control]. *Revue d’ Epidémiologie et de Médecine Préventive [Journal of Epidemiology and Preventive Medicine]*. 2003; 1: 15-22.

[15] Bobenko AI. *Geometry II: Discrete Differential Geometry.* TU Berlin, Germany; 2015.