

TEACHING SPLINE APPROXIMATION TECHNIQUES USING MAPLE¹

Mihály Klincsik, Pécs (Hungary)

Abstract: Using Computer Algebra Systems (CAS) - such as MAPLE- in teaching and learning mathematical concepts is a great challenge both from a didactical and a scientific point of view. We have to rewrite our traditional paper based teaching materials for interactive and living electronic worksheets. Only few statements and principles have to be acquired by the learner and the teacher from the CAS and after they can visualise, make animations, modify quickly the program data, perform symbolic and numeric calculations step by step and in the whole, and verify deductions on their own. The author prepared Maple worksheets for teaching different types of function approximating techniques, such as interpolation-, least square-, spline and uniform approximation methods for post-graduate mechanical engineering students. In this paper we want to demonstrate how can we keep and improve the famous problem solving principles and rules given by G. Pólya and R. Descartes (Pólya 1962), when we use the capabilities of CAS. The education principles: active learning, motivations and the successive phases are getting new meaning in the CAS. Our examples are always concerning with spline functions. Handling the formulas, calculating values and giving proofs are always in the form of Maple statements.

Kurzreferat:

Benutzung von CAS im Lehr- und Lernprozess ist eine große didaktische und wissenschaftliche Herausforderung. Alle unsere didaktischen Prinzipien sollen in Hinsicht des Gebrauches von CAS neu durchgedacht werden. Auf der anderen Seite braucht man nur einige CAS Hilfsmittel zu erlernen, um ein CAS als ein leistungsfähiges Werkzeug zur Unterstützung der Modellbildung, zur Visualisierung und bei numerischen und symbolischen Rechnungen einsetzen zu können.

Der Verfasser stellte Maple Arbeitsblätter für den Unterricht der verschiedenen Verfahren der Approximationstheorie zusammen. Diese Lehrmaterialien werden im Fernunterricht verwendet. Es wird versucht zu demonstrieren, wie erfolgreich Maple CAS eingesetzt werden kann, um die didaktischen Prinzipien von G. Polya und R. Descartes im Unterricht zu verwirklichen. Als Beispiel wählen wir immer Beispiele aus der Splinetheorie. Zur Bearbeitung benutzen wir immer Maple .

ZDM -Classification: D50, N50, U50

1. Introduction

Mass education, re-training courses and learning beside job enforce developments and adaptations of efficient and flexible distance educational methods in teaching math, too. Realising these facts we have to rewrite our traditional math text-base lecture notes, applying the knowledge-based **Computer Algebraic System (CAS)** and we have to give chances for the users that they can access these materials via computer **network**.

Choosing one of the post-graduate courses for the mechanical engineering students at the Engineering Faculty of Pécs University we have prepared **Maple worksheets**, teaching of different types of function approximation techniques: such as **interpolation-, least squares-, spline** and **uniform** approximation and the **numerical approximation** of solution of **ordinary differential equations**. Assuming college level knowledge of math calculus and linear algebra, the Maple language and the usage of the **Internet** for teaching purposes were new for the students.

Before the availability of CAS computer program languages (FORTRAN, Pascal, Basic, and C) were the only essential tools for teaching and demonstrating different types of function approximation techniques. Large collections of scientific and statistical subroutine packages were made and the users could solve his/her problems calling the routines with different actual parameters. In the teaching process the usage and analysis of these procedures helped to understand the algorithms and the re-programming activity supported the remembering process.

What kind of new activities and results can we achieve with the application of CAS and the Internet in teaching of this special mathematical field: **function approximation**? On the basis of our 5-year practice we can answer: using both CAS and Internet together improves dramatically the attitudes both of the teachers and the students in regard to the following areas:

- the learner is an active participant of the learning process;
- it improves our abilities to analyse the problem in analytical, graphical, symbolical and numerical forms;
- the user can focus on the essential steps of the problem solving, and don't waste the time in the details;
- the proofs and the deductions can be creating throughout programming statements without mistakes;
- the teaching materials and the course information can be easily modified and accessed;
- the public information about the courses (learning materials, date of exams, results) are always up-to-date.

2. Demonstration and improvement of the learning model given by G. Pólya using the capabilities of CAS

2.1 Use the potential capabilities offered by the CAS to motivate and stimulate the users for activities

When an engineer meets with the concept of the spline function approximation technique for the first time he/she wants to know the answers for the following questions:

¹ Modified version of a paper presented at the International Symposium, Anniversary of Pollack Mihály Engineering Faculty, May 31-June 1, 2002, Pécs, Hungary

- What is the new in this concept in comparison with the interpolation?
- Why it is important besides the interpolation?
- How can we calculate and draw the function, when we apply it for an engineering problem?

The **basic construction** and the **definition** of the spline are very simple: a **spline function** is such a function that is consisting of polynomial pieces on different subintervals and these polynomials are joining together with certain smoothness (differentiability and continuity) conditions. The approximation means that the spline function passes through the given points.

We can show **immediately** the spline approximation functions for the students according to the different degrees of polynomials using the built-in Maple procedure **'spline'**. The calling sequence of the spline procedure is

spline(X, Y, z, d),

where the parameters **X** and **Y** are the given vectors or lists of the abscissas and the function values respectively, **z** is the name of the independent variable for the obtained spline function and **d** is the degree of the polynomials.

The returning values of this procedure form a **piecewise polynomial** function with degree **d**, giving the coefficients of the different polynomials on different subintervals.

As an easy example, we choose the well known function $y = \sin(x)$ to approximate it on the interval $[0,6]$ in integer abscissas using the degrees 1, 2 and 3 of the polynomial functions. Consider the simple Maple commands to generate the linear spline approximation function for the sin function

```
>T:=seq(i,i=0..6);
>Y:=map(evalf@sin,T);
>Linear:=spline(T,Y,x,1);
```

$$Linear := \begin{cases} .8414709848 x & x < 1 \\ .7736445428 + .0678264420 x & x < 2 \\ 2.445652263 - .7681774182 x & x < 3 \\ 2.834887518 - .8979225033 x & x < 4 \\ .05168462150 - .2021217792 x & x < 5 \\ -4.356468161 + .6795087773 x & otherwise \end{cases}$$

Immediately we can **plot** the graph of this function and thus the students have both **analytical** and **visual** concepts about the linear splines:

```
>plot([Linear,sin(x)],x=T[1]..T[nops(T)],color=[red,blue]);
```

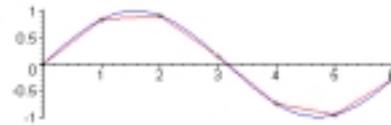


Fig. 1: Approximation of the $\sin(x)$ function using linear spline on $[0,6]$ interval in integer abscissas

The picture (see Fig. 1.) shows that the linear spline connects successive points with straight lines, calling it polygonal line. The straight lines can't follow the bends of the sin function, thus the error (maximal difference between the linear spline and the sin function on the whole interval) is too large.

How can we modify the idea of this linear spline function to get a better approximation? The answer is giving by the students immediately with **increasing the degree** of the piecewise polynomials. Now we plot the graphs of the second and third degree piecewise polynomial spline functions - generate them by the Maple commands (see Fig. 2.)

```
spline(T,Y,x,2); and spline(T,Y,x,3);
```

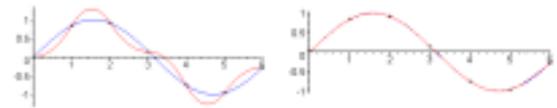


Fig. 2: Quadratic and cubic spline approximations of the $\sin(x)$ function on the interval $[0,6]$ in integer abscissas

Investigating the graph of the **quadratic spline** function we can observe that this spline takes some unreasonable wide bends and therefore the result is **worse** than it is in the linear case, the error is much larger. This oscillation behaviour characterises the interpolation functions, when the degree (and so the number of the points) of the approximation polynomial is too high.

The **natural cubic spline** (see Fig. 2.) is so smooth and approximates the sin function so good that we can't distinguish the sin function from their spline function on the basis of the picture. But we can say similar things about the **interpolation** polynomial (which degree is 6) in this situation, too.

The students can approximate other functions by their own, knowing only these few statements from MAPLE. Such a way the students become **active participants** of these techniques from the first moment. Using these techniques and make some computer experiments they can answer for these questions:

- Whether the derivatives of a cubic spline can result another spline function with smaller degree polynomials?
- What kind of function operations can be performing on the set of cubic spline functions to remaining in this set?

2.2 Motivate the students by giving the origin and the applications of the cubic spline

The **cubic spline** gives the best approximation among the previously mentioned three splines, therefore the engineering students want to know more about their applications and origin.

Planning the profiles of a ship, aeroplane wings and the car's spoilers are motivated the application of the cubic spline approximation. The methods are based on the results of the wind tunnel experiments. The engineers draw the points (call them knots) getting from the experiments to a table and then were fitting a flexible thin wooden slat among these points. The thin slat was forming a banding curved shape, joining the knots on the table and can be draw the contour of this shape. (See Fig. 3.)

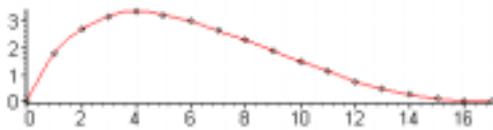


Fig. 3: A streamline wing profile approximating it by cubic spline

The designer knew that a bar shape could describe by a cubic polynomial, when it is loaded at their endpoints. But in this situation one cubic polynomial function can't yield so many kinds of shapes and so many forces act for the thin slat, not only two. Therefore the idea arises: can we approximate these shapes by piecewise cubic polynomials. When we analyze the problem by mathematical tools we should keep in mind the method of G. Pólya teaching the problem in successive phases from the easy toward the more complex ones.

3. Analyse the problem solvability and the solution of the cubic spline according to the principles of R. Descartes

Now let us consider in simple form of the formulations of the natural cubic spline by mathematical formulas. We want to follow the basic principles and rules suggested by R. Descartes, how can we formulate and solve generally a problem by using of mathematical tools. (Pólya 1962) We will denote the **Rules of Descartes** by **RD**.

3.1 Let us trace back the well understood and posed problem to determining certain unknown quantities. (RD 1.)

We are searching for the natural cubic spline function **S(x)** at the tabulated data

Abcissas	t_1	t_2	t_3	...	t_{n-1}	t_n
Function values	y_1	y_2	y_3	...	y_{n-1}	y_n

where the n abcissas are different from each other and ordered by increasing.

The function **S(x)** satisfies the following previously mentioned requirements, prescriptions

(UQ) **S(x)** equals to the third degree polynomial function

$$S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \quad (i = 1, 2, \dots, (n-1))$$

on the subinterval $[t_i, t_{i+1}]$ (on different subintervals the polynomials are different).

Now we have $4 \cdot (n-1)$ **Unknown Quantities (UQ)** which are the coefficients of the cubic polynomial pieces. We have to specify therefore enough conditions to determine these unknown quantities uniquely.

Assumptions for the spline **S** are the following

(A1) the spline function passes through the tabulated points

$$S(t_i) = y_i \quad (i = 1, 2, \dots, n),$$

(A2-4) the function **S**, the first and the second derivative of **S**

$$S(x), S'(x) = \frac{dS(x)}{dx}, S''(x) = \frac{d^2 S(x)}{dx^2}$$

are continuous on the whole interval $[t_1, t_n]$.

Let us **calculate** how many (unknown) coefficients we have and how many conditions, equations we have to determine the cubic spline function **S**.

Now we have **4 (n-1) unknown coefficients** on bases of the forms **(UQ)**, because of there are $(n-1)$ different subintervals and on each subinterval there are 4 coefficients of the searching cubic polynomial function.

But we have

- **n** equations using the condition **(A1)**, because of the function passing through the given n knots (points);
- **3(n-2)** equations using the continuity of the function **S(x)** and the derivatives

$$S'(x) = \frac{dS(x)}{dx}, S''(x) = \frac{d^2 S(x)}{dx^2} \quad \text{only at the}$$

interior abcissas $x = t_2, t_3, \dots, t_{n-1}$.

Thus the total sum of the equations are $n+3(n-2) = 4n-6$, which is less than by 2 from the number $(4n-4)$, the number of the unknown coefficients. Therefore we can **prescribe** 2 further new conditions for the function **S(x)** - which we have to choose independently from the conditions **(A1)**, **(A2)**, **(A3)** and **(A4)** ones - in order to use all degrees of the freedom.

In the case of **natural cubic spline** these 2 conditions are imposed for the boundary values of the second derivative of **S**

$$(A5) \quad S''(t_1) = S''(t_n) = 0.$$

3.2 Let us consider that the problem is solved and transform the quantities and relations into more convenient form to get fewer unknowns and equations. (RD 2.)

Because of $S_i(x)$ is a cubic polynomial function, therefore their second derivative is linear function. Write the linear interpolation for the function $S_i''(x)$ in the following form using the Maple

$$\frac{\partial^2}{\partial x^2} S_i(x) = \frac{z_{i+1}(x-t_i)}{t_{i+1}-t_i} + \frac{z_i(t_{i+1}-x)}{t_{i+1}-t_i}$$

From the assumption (A4) the function $S''(x)$ is a continuous piecewise linear function, thus the values $z_i = S''(t_i)$ are well defined. The values z_2, z_3, \dots, z_{n-1} are unknown, but $z_1 = z_n = 0$ are known because of the boundary conditions (A5).

Now we can construct the natural cubic spline $S(x)$ through the values z_i ($i = 2, 3, 4, \dots, n-1$).

But the number of these unknown quantities are only $(n-2)$, comparing with the number $4(n-1)$ in the case of (RD 1.). Where we loss information? In this formulation now we have already used $(n-2)$ assumptions from the continuity of second derivative of S and the 2 boundary conditions (A5). Therefore the remaining conditions are only $4(n-1) - n = 3n - 4$. We will show that we can construct $S(x)$ in this way.

If we integrate the above relation twice in the Maple, then we can obtain $S_i(x)$ itself in the form

$$S_i(x) = -\frac{1}{6} \frac{z_{i+1}(x-t_i)^3}{-t_{i+1}+t_i} + \frac{1}{6} \frac{z_i(-t_{i+1}+x)^3}{-t_{i+1}+t_i} + C1(x-t_i) + C2(t_{i+1}-x)$$

If we use of the assumption (A1) and the continuity assumption (A2) we have got the integration constants $C1$ and $C2$ in the forms:

$$C1 := \frac{y_{i+1}}{h_i} - \frac{1}{6} z_{i+1} h_i$$

$$C2 := \frac{y_i}{h_i} - \frac{1}{6} z_i h_i$$

$$t_{i+1} - t_i = h_i$$

Therefore we can obtain the cubic spline in the following compact form

$$S_i(x) = \frac{1}{6} \frac{z_{i+1}(x-t_i)^3}{h_i} - \frac{1}{6} \frac{z_i(-t_{i+1}+x)^3}{h_i} - \frac{1}{6} \frac{(-6y_{i+1}+z_{i+1}h_i^2)(x-t_i)}{h_i} + \frac{1}{6} \frac{(-6y_i+z_ih_i^2)(-t_{i+1}+x)}{h_i}$$

3.3 Use the remainder assumptions to get system of equations for the unknown quantities (RD 3.)

Now let us handle separately those of all conditions from which we can determine the same quantity in two different ways. Summarise that the assumptions (A1), (A2), (A4) and (A5) we have already used to obtain the form of functions $S_i(x)$. Using now the continuity of the first derivative S' - assumption (A3) - we have the conditions

$$S'_{i-1}(t_i) = S'_i(t_i) \quad (i = 2, 3, \dots, n-1).$$

Formally by the Maple we can derive the formulas obtained for S_i and S_{i-1} , and substitute the value t_i into these functions symbolically. Then we conclude from this equation the following relations

for the index $i = 2, 3, \dots, (n-1)$. This is a system of $(n-2)$

$$h_{i-1} z_{i-1} + (2h_{i-1} + 2h_i) z_i + z_{i+1} h_i = \frac{6y_{i+1} - 6y_i}{h_i} + \frac{-6y_i + 6y_{i-1}}{h_{i-1}}$$

linear equations with the same number of unknowns is appropriate for determining the z -values. Thus we can get the spline function S solving a smaller number of equations as we mentioned at the beginning of 3.1.

3.4 Let us reduce the system of equations into only one equation. (RD 4.)

We obtained a special structure of the system of linear equations - to determine the z -value of natural cubic spline S - whose coefficient matrix is tridiagonal. All elements of a tridiagonal matrix are zeros except possibly for those situated on the main diagonal, the subdiagonal, and the superdiagonal, i.e. over and upper band near the diagonal.

In order to obtain the solution of these linear equations

we rewrite our tridiagonal system into the following matrix form

$$\begin{bmatrix} b_1 & c_1 & 0 & 0 & \dots & 0 \\ a_2 & b_2 & c_2 & 0 & \dots & 0 \\ 0 & a_3 & b_3 & c_3 & \dots & 0 \\ 0 & 0 & a_4 & b_4 & c_4 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & 0 & 0 & 0 & a_n & b_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_{n-1} \\ r_n \end{bmatrix}$$

or equivalently in the equations form

$$(1) \quad a_i \cdot x_{i-1} + b_i \cdot x_i + c_i x_{i+1} = r_i \quad (i = 1, 2, \dots, n)$$

where $a_1 = c_n = 0$.

We can consider these equations as a second order linear **recurrent relation**, because of x_{i-1}, x_i, x_{i+1} appear in the equation. From this point of view we handle this problem to prepare a procedure solving such a special band structure linear system.

From the first equation of the system (1) we can express x_1 with expression x_2 in linear form. Substitute it into the second equation then we have got a same expression for x_2 with x_3 . We can follow these

substitutions and have got from the (n-1)-th equation an expression for x_{n-1} with using x_n . Finally we can solve the last equation exactly for x_n , when we substitute the expression for x_{n-1} into the last equation. We mention that this algorithm is working, because of the diagonal dominant of the band matrix. Now the proof of the solvability of the cubic spline function is completed theoretically and the rules of Descartes were adjusted to this problem.

4. Next phase is going from the top-down method toward to the bottom-up method using the CAS

The above-mentioned mathematical description to obtain the solution of the cubic spline problem was a typical example for planning a lecture or theory by the use of **top-down** method. This means that in the beginning we study the entire problem on the whole and after we divide it into smaller pieces. Then we try to handle these pieces, boxes of the problem in a similar way, that the entire problem. In this way we try to get acquainted with the inside of the **black boxes**, i.e. the black boxes successively become whiter and whiter (Sárvári 2002). These mathematical descriptions are revealing the veil about the covered important details step by step. The solution is appearing from the anonymity darkness. As we have seen the knowledge become more and more deep on the following way

- (a) We know only the definition of the spline;
- (b) We have an analytical description;
- (c) We transform the problem into new variables;
- (d) We set up equations for the unknowns;
- (e) We solve the linear equations;

If we don't use the CAS then the learning and teaching process is terminated here. But this process will be turning back when we try to use the CAS program capabilities.

4.1 The bottom-up method: preparation of the CAS program from the mathematical algorithm

When the students are familiar with the steps of the solution, then we can prepare the detailed algorithms and the Maple programs of these steps. In this active learning phase of the teaching process we can get certainly new information about the solvability and the methods of the problems, but the real aims are not these. When we use the CAS to prepare procedures for solving a more complex problem we built our knowledge into these procedures.

The descriptions of the problems by the way of CAS are hiding the unimportant information about the user. The solution is remaining in the anonymity darkness; we can see only the inputs and the outputs, the results. This modular process is the reverse as we mentioned previously and we build from the bottom direction to the upward:

- (a) Write a procedure for solving a linear system of equations with **tridiagonal matrix**;
- (b) Write a procedure **to set up the equations** from the given data of the abscissa and the ordinates and **calculate the z-values** using the previous procedure;
- (c) Write a procedure to **create the piecewise cubic polynomial functions** using the calculated z-values output from the previous procedure;

The aim of this learning process is fixing the knowledge about the steps in the mind. At the end of this process we have got black boxes from our thorough knowledge to build in.

5. Next is the experimenting phase for discovering the well-known minimization properties of the natural cubic spline

We can verify and prove the well-known minimization properties of the natural cubic spline using the Maple. Beginning with numerical computation, determine the interpolation polynomial for the function $y = e^x$ on the interval $[0,1]$ dividing into 5 equal pieces. Then we can calculate the integral of the second power of the second derivative of the 5 degrees interpolation polynomial over the interval $[0,1]$ with the Maple command:

```
>Int(Diff(interpol,z$2)^2,z=a..b)=
evalf(int(diff(interpol,z$2)^2,z=a..b)
);
```

The result is the following

$$\int_0^1 \left(\frac{\partial^2}{\partial x^2} (.01385411458 z^5 + .03486687501 z^4 + .1704093854 z^3 + .4990689250 z^2 + 1.000082528 z + 1.) \right)^2 dz = 3.193847417$$

In the next step we determine the natural cubic spline approximates the same function in the same points, and calculate the integral of the second power of the second derivative of the piecewise polynomial over the interval [0,1] with the Maple command

```
>Int(Diff(spline3,z$2)^2,z=a..b)=evalf
(int(diff(spline3,z$2)^2,z=a..b));
```

And now the result is

$$\int_0^1 \left(\frac{\partial^2}{\partial x^2} \begin{cases} 1. + 1.057966300 z + 1.226187259 z^3 & z < \frac{1}{5} \\ 1.009837807 + .9103992033 z + .7378354820 z^2 - .0035385444 z^3 & z < \frac{2}{5} \\ 1.000357619 + .9815006106 z + .5600819647 z^2 + .1445893868 z^3 & z < \frac{3}{5} \\ .7978415002 + 1.994081206 z - 1.127552361 z^2 + 1.082164012 z^3 & z < \frac{4}{5} \\ 2.606003356 - 4.786525755 z + 7.348206341 z^2 - 2.449402114 z^3 & otherwise \end{cases} \right)^2 dz = 2.712422014$$

We can see that the value of the integral is less in the case of the spline function. The students can take some other quick experiments with additional functions not only with the exponential ones to make sure that this phenomenon is not an accidental event.

Now the student explored the famous behavior of the natural cubic spline function. After these events we present the theorem and the proof with exact formulation for which we are saying that the cubic splines are the best function to employ for curve fitting.

Theorem. Let the function $S(x)$ is the natural cubic spline and $y(x)$ is an arbitrary twice continuously differentiable function for which the conditions $y(t_i) = S(t_i)$ are fulfilled ($i = 1, 2, \dots, n$). Then the inequality

$$\int_a^b \left(\frac{\partial^2}{\partial x^2} S(x) \right)^2 dx \leq \int_a^b \left(\frac{\partial^2}{\partial x^2} y(x) \right)^2 dx$$

is valid.

References

- [1] Ward Cheney, David Kincaid, Numerical Mathematics and Computing, Brooks/Cole Publishing Company Monterey, California, (1980).
- [2] J. LI. Morris, Computational Methods in Elementary Numerical Analysis, John Wiley & Sons, (1983).
- [3] G. Pólya, Mathematical discovery, on understanding, learning and teaching problem solving, John Wiley & Sons, (1962).
- [4] Cs. Sárvári, Zur Möglichkeiten der flexiblen, transferreichen Erlernung mit Hilfe von CAS, p. 431-434. Beitrage zum Mathematikunterricht, 2002.

Autor

Mihály Klincsik, Ph.D., University of Pécs, Dept. of Mathematics of Pollack Mihály Engineering Faculty, H-7624 Pécs, Boszorkány út 2.

E-mail: klincsik@witch.pmmf.hu