Searching for the best Points of interpolation using swarm intelligence techniques

Djerou L., Khelil N., Zerarka A. and Batouche M.

1. Introduction

If the values of a function, \( f(x) \), are known for a finite set of \( x \) values in a given interval, then a polynomial which takes on the same values at these \( x \) values offers a particularly simple analytic approximation to \( f(x) \) throughout the interval. This approximating technique is called polynomial interpolation. Its effectiveness depends on the smoothness of the function being sampled (if the function is unknown, some hypothetical smoothness must be chosen), on the number and choice of points at which the function is sampled.

In practice interpolating polynomials with degrees greater than about 10 are rarely used. One of the major problems with polynomials of high degree is that they tend to oscillate wildly. This is clear if they have many roots in the interpolation interval. For example, a degree 10 polynomial with 10 real roots must cross the x-axis 10 times. Thus, it would not be suitable for interpolating a monotone decreasing or increasing function on such an interval.

In this chapter we explore the advantage of using the Particle Swarm Optimization (PSO) interpolation nodes. Our goal is to show that the PSO nodes can approximate functions with much less error than Chebyshev nodes.

This chapter is organized as follows. In Section 2, we shall present the interpolation polynomial in the Lagrange form. Section 3 examines the Runge’s phenomenon; which illustrates the error that can occur when constructing a polynomial interpolant of high degree. Section 4 gives an overview of modern heuristic optimization techniques, including fundamentals of computational intelligence for PSO. We calculate in Subsection 4.2 the best interpolating points generated by PSO algorithm. We make in section 5, a comparison of interpolation methods. The comments and conclusion are made in Section 6.

2. Introduction to the Lagrange interpolation

If \( x_0, x_1, ..., x_n \) are distinct real numbers, then for arbitrary values \( y_0, y_1, ..., y_n \) there is a unique polynomial \( p_n \) of degree at most \( n \) such that \( p_n(x_i) = y_i \) \((0 \leq i \leq n)\) (David Kincaid & Ward Cheney, 2002).

The Lagrange form looks as follows:
\[ p_n(x) = y_0l_0(x) + \ldots + y_nl_n(x) = \sum_{i=0}^{n} y_il_i(x) \]  

(1)

Such that coordinal functions can be expressed in the following

\[ l_i(x) = \prod_{j=0, j \neq i}^{n} \frac{(x - x_j)}{(x_i - x_j)} \]  

(2)


are coordinal polynomials that satisfy

\[ l_i(x_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \]  

(3)

The Lagrange form gives an error term of the form

\[ E_n(x) = f(x) - p_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \phi_n(x) \]  

(4)

Where

\[ \phi_n(x) = \prod_{i=0}^{n} (x - x_i) \]  

(5)

If we examine the error formula for polynomial interpolation over an interval \([a, b]\) we see that as we change the interpolation points, we change also the locations \(c\) where the derivative is evaluated; thus that part in the error also changes, and that change is a "black hole" to us: we never know what the correct value of \(c\) is, but only that \(c\) is somewhere in the interval \([a, b]\). Since we wish to use the interpolating polynomial to approximate the Equation (4) cannot be used, of course, to calculate the exact value of the error \(f - P_n\), since \(c\), as a function of \(x\) is, in general, not known. (An exception occurs when the \((n + 1)st\) derivative off is constant). And so we are likely to reduce the error by selecting interpolation points \(x_0, x_1, \ldots, x_n\) so as to minimize the maximum value of product \(\phi_n(x)\).

The most natural idea is to choose them regularly distribute in \([a, b]\).

3. Introduction to the Runge phenomenon and to Chebyshev approximations

3.1 Runge phenomenon

If \(x_k\) are chosen to be the points \(x_k = a + k \frac{b-a}{n} \) \((k = 0, \ldots, n)\) (means that are equally spaced at a distance \(2n + 1\) apart), then the interpolating polynomial \(p_n(x)\) need not to converge uniformly on \([a, b]\) as \(n \rightarrow \infty\) for the function \(f(x)\).
This phenomenon is known as the Runge phenomenon (RP) and it can be illustrated with the Runge's "bell" function on the interval [-5, 5] (Fig. 1).

Figure 1. solid blue line: present Runge’s “bell” function. dots red line: present the polynomial approximation based on equally 11 spaced nodes

3.2 Chebyshev Nodes
The standard remedy against the RP is Chebyshev -type clustering of nodes towards the end of the interval (Fig. 3).

Figure 2. Chebyshev Point Distribution.
To do this, conceptually, we would like to take many points near the endpoints of the interval and few near the middle. The point distribution that minimizes the maximum value of product $\phi_n(x)$ is called the Chebyshev distribution, as shown in (Fig. 2). In the Chebyshev distribution, we proceed as follows:
1. Draw the semicircle on [a, b].
2. To sample $n + 1$ points, place $n$ equidistant partitions on the arc.
3. Project each partition onto the $x$-axis: for $j = 0, 1 \ldots n$

$$x_j = \frac{a + b}{2} + \frac{b - a}{2} \cos \left( j \frac{\pi}{n} \right) \text{ for } j = 0, 1 \ldots n$$

The nodes $x_i$ that will be used in our approximation are:

<table>
<thead>
<tr>
<th>Chebyshev nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5.0000</td>
</tr>
<tr>
<td>-4.2900</td>
</tr>
<tr>
<td>-4.0251</td>
</tr>
<tr>
<td>-2.6500</td>
</tr>
<tr>
<td>-1.4000</td>
</tr>
<tr>
<td>0.0000</td>
</tr>
<tr>
<td>1.4000</td>
</tr>
<tr>
<td>2.6500</td>
</tr>
<tr>
<td>4.0451</td>
</tr>
<tr>
<td>4.2900</td>
</tr>
<tr>
<td>5.0000</td>
</tr>
</tbody>
</table>

Figure 3. solid blue line: present Runge’s “bell” function. dots red line: present the polynomial approximation based on 11 Chebyshev nodes.

In this study, we have made some numerical computations using the particle swarm optimization to investigate the best interpolating points and we are showing that PSO nodes provide smaller approximation error than Chebyshev nodes.

4. Particle swarm optimization

4.1 Overview and strategy of particle swarm optimization

Recently, a new stochastic algorithm has appeared, namely ‘particle swarm optimization’ PSO. The term ‘particle’ means any natural agent that describes the ‘swarm’ behavior. The PSO model is a particle simulation concept, and was first proposed by Eberhart and Kennedy (Eberhart, R.C. et al. 1995). Based upon a mathematical description of the social
behavior of swarms, it has been shown that this algorithm can be efficiently generated to find good solutions to a certain number of complicated situations such as, for instance, the static optimization problems, the topological optimization and others (Parsopoulos, K.E. et al., 2001a); (Parsopoulos, K.E. et al. 2001b); (Fourie, P.C. et al., 2000); (Fourie, P.C. et al., 2001). Since then, several variants of the PSO have been developed (Eberhart, R.C. et al 1996); (Kennedy, J. et al., 1998); (Kennedy, J. et al., 2001); (Shi, Y.H. et al. 2001); (Shi, Y. et al. 1998a.); (Shi, Y.H. et al., 1998b); (Clerc, M. 1999 ). It has been shown that the question of convergence of the PSO algorithm is implicitly guaranteed if the parameters are adequately selected (Eberhart, R.C. et al.1998); (Cristian, T.I. 2003). Several kinds of problems solving start with computer simulations in order to find and analyze the solutions which do not exist analytically or specifically have been proven to be theoretically intractable.

The particle swarm treatment supposes a population of individuals designed as real valued vectors - particles, and some iterative sequences of their domain of adaptation must be established. It is assumed that these individuals have a social behavior, which implies that the ability of social conditions, for instance, the interaction with the neighborhood, is an important process in successfully finding good solutions to a given problem.

The strategy of the PSO algorithm is summarized as follows: We assume that each agent (particle) \( i \) can be represented in a \( N \) dimension space by its current position \( x_i = (x_{i1}, x_{i2}, \ldots, x_{iN}) \) and its corresponding velocity. Also a memory of its personal (previous) best position is represented by, \( p = (p_{i1}, p_{i2}, \ldots, p_{iN}) \) called (pbest), the subscript \( i \) range from 1 to \( s \), where \( s \) indicates the size of the swarm. Commonly, each particle localizes its best value so far (pbest) and its position and consequently identifies its best value in the group (swarm), called also (sbest) among the set of values (pbest).

The velocity and position are updated as

\[ v_i^{k+1} = w v_i^k + c_1 r_1 (p_{best}^k - x_i^k) + c_2 r_2 (s_{best}^k - x_i^k) \]  

(7)

\[ x_i^{k+1} = v_i^{k+1} + x_i^k \]  

(8)

where are the position and the velocity vector of particle \( i \) respectively at iteration \( k + 1 \), \( c_1 \) and \( c_2 \) are acceleration coefficients for each term exclusively situated in the range of 2–4, \( w_{ij} \) is the inertia weight with its value that ranges from 0.9 to 1.2, whereas \( r_1 \), \( r_2 \) are uniform random numbers between zero and one. For more details, the double subscript in the relations (7) and (8) means that the first subscript is for the particle \( i \) and the second one is for the dimension \( j \). The role of a suitable choice of the inertia weight \( w_{ij} \) is important in the success of the PSO. In the general case, it can be initially set equal to its maximum value, and progressively we decrease it if the better solution is not reached. Too often, in the relation (7), \( w_{ij} \) is replaced by \( w_{ij} / \sigma \), where \( \sigma \) denotes the constriction factor that
controls the velocity of the particles. This algorithm is successively accomplished with the following steps (Zerarka, A. et al., 2006):

1. Set the values of the dimension space N and the size s of the swarm (s can be taken randomly).
2. Initialize the iteration number k (in the general case is set equal to zero).
3. Evaluate for each agent, the velocity vector using its memory and equation (7), where pbest and sbest can be modified.
4. Each agent must be updated by applying its velocity vector and its previous position using equation [8].
5. Repeat the above step (3, 4 and 5) until a convergence criterion is reached.

The practical part of using PSO procedure will be examined in the following section, where we’ll interpolate Runge’s “bell”, with two manners; using Chebyshev interpolation approach and PSO approach, all while doing a comparison.

4.2 PSO distribution

So the problem is the choice of the points of interpolation so that quantity $\phi_n(x)$ deviates from zero on [a, b] the least possible.

Particle Swarm Optimization was used to find the global minimum of the maximum value of product $\phi_n(x)$, where very $x$ is represented as a particle in the swarm.

The PSO parameter values that were used are given in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size</td>
<td>20</td>
</tr>
<tr>
<td>Number of iterations</td>
<td>500</td>
</tr>
<tr>
<td>C1 and C2</td>
<td>0.5</td>
</tr>
<tr>
<td>Inertial Weight</td>
<td>1.2 to 0.4</td>
</tr>
<tr>
<td>Desired Accuracy</td>
<td>10-5</td>
</tr>
</tbody>
</table>

Table 1. Particle Swarm Parameter Setting used in the present study

The best interpolating points $x$ generated by PSO algorithm for polynomial of degree 5 and 10 respectively for example are:

<table>
<thead>
<tr>
<th>Chebyshev</th>
<th>Points generated with PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5.0000</td>
<td>-5.0000</td>
</tr>
<tr>
<td>-3.9355</td>
<td>-4.0451</td>
</tr>
<tr>
<td>-2.9041</td>
<td>-1.5451</td>
</tr>
<tr>
<td>0.9000</td>
<td>1.5451</td>
</tr>
<tr>
<td>3.9355</td>
<td>4.0451</td>
</tr>
<tr>
<td>5.0000</td>
<td>5.0000</td>
</tr>
</tbody>
</table>

Table 2 Polynomial of degree 5
Searching for the best Points of interpolation using swarm intelligence techniques

<table>
<thead>
<tr>
<th>Chebyshev</th>
<th>Points generated with PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5.0000</td>
<td>-5.0000</td>
</tr>
<tr>
<td>-4.2900</td>
<td>-4.7533</td>
</tr>
<tr>
<td>-4.0251</td>
<td>-4.0451</td>
</tr>
<tr>
<td>-2.6500</td>
<td>-2.9389</td>
</tr>
<tr>
<td>-1.4000</td>
<td>-1.5451</td>
</tr>
<tr>
<td>0.0000</td>
<td>-0.0000</td>
</tr>
<tr>
<td>1.4000</td>
<td>1.5451</td>
</tr>
<tr>
<td>2.6500</td>
<td>2.9389</td>
</tr>
<tr>
<td>4.0451</td>
<td>4.0451</td>
</tr>
<tr>
<td>4.2900</td>
<td>4.7553</td>
</tr>
<tr>
<td>5.0000</td>
<td>5.0000</td>
</tr>
</tbody>
</table>

Table 3. Polynomial of degree 10

5. Comparison of interpolation methods

How big an effect can the selection of points have? Fig. 4 and Fig. 5 shows Runge's "bell" function interpolated over [-5, 5] using equidistant points, points selected from the Chebyshev distribution, and a new method called PSO. The polynomial interpolation using Chebyshev points does a much better job than the interpolation using equidistant points, but neither does as well as the PSO method.

![Comparison of interpolation polynomials](image1)

![Error Plot](image2)

Figure 4. Comparison of interpolation polynomials for equidistant and Chebyshev sample points
Comparing Fig. 4, we see that the maximum deviation of the Chebyshev polynomial from the true function is considerably less than that of Lagrange polynomial with equidistant nodes. It can also be seen that increasing the number of the Chebyshev nodes—or, equivalently, increasing the degree of Chebyshev polynomial—makes a substantial contribution towards reducing the approximation error.

Comparing Fig. 5, we see that the maximum deviation of the PSO polynomial from the true function is considerably less than that of Chebyshev polynomial nodes. It can also be seen that increasing the number of the PSO nodes—or, equivalently, increasing the degree of PSO polynomial—makes a substantial contribution towards reducing the approximation error.

Figure 5. Comparison of interpolation polynomials for PSO and Chebyshev sample points

In this study we take as measure of the error of approximation the greatest vertical distance between the graph of the function and that of the interpolating polynomial over the entire interval under consideration (Fig. 4 and Fig. 5).

The calculation of error gives

<table>
<thead>
<tr>
<th>Degree</th>
<th>Error points equidistant</th>
<th>Error points Chebychev</th>
<th>Error points PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.4327</td>
<td>0.6386</td>
<td>0.5025</td>
</tr>
<tr>
<td>10</td>
<td>1.9156</td>
<td>0.1320</td>
<td>0.1076</td>
</tr>
<tr>
<td>15</td>
<td>2.0990</td>
<td>0.0993</td>
<td>0.0704</td>
</tr>
<tr>
<td>20</td>
<td>58.5855</td>
<td>0.0177</td>
<td>0.0131</td>
</tr>
</tbody>
</table>

Table 2. The error
6. Conclusion

The particle swarm optimization is used to investigate the best interpolating points. Some good results are obtained by using the specific PSO approach. It is now known that the PSO scheme is powerful, and easier to apply specially for this type of problems. Also, the PSO method can be used directly and in a straightforward manner. The performance of the scheme shows that the method is reliable and effective.

7. References


Particle swarm optimization (PSO) is a population based stochastic optimization technique influenced by the social behavior of bird flocking or fish schooling. PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles. This book represents the contributions of the top researchers in this field and will serve as a valuable tool for professionals in this interdisciplinary field.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: