

Grzegorz Lenda\*

## The Application of Least-Squares Method for Approximating the Surfaces of Engineering Structures

### 1. Introduction

When the shapes of engineering objects are examined, very often there is a need for changing the point set of data acquired during survey and converting it to a mathematical, continuous model representing the specific object. The author's experience shows that interpolation or approximation by means of splines is the best solution. The splines, as compared to classical methods of LSM approximation, e.g. with the use of second-degree surfaces, are characteristic of several very important advantages. The most essential of them is the possibility of acquiring a continuous description of the structure shape, containing its local deformations, unlike in classical methods where the ideal object (sphere, cylinder, plane) is obtained as a result of approximation, and information regarding deformations may be acquired only through comparison of the point set obtained from the survey with this object. Such an approach significantly hinders performing analyses in the course of periodical observations and inspections of building structures, particularly when surveys are performed each time in different places of this building structure. Yet another advantage of splines involves the absence of the necessity of knowing the type of mathematical function, in accordance with which the structure has been erected as the spline is in each case created by polynomials joined with continuity  $C^2$ . Where the classical approach is considered and the approximating surface inappropriately selected, major errors of matching are inevitable. The locality of the description is the last advantage of splines. The particular points measured exert influence not on the shape of the entire surface modelled, but only on its minor part, so there is no risk that any local deformations or measuring error will have impact on the entire structure.

---

\* Department of Mining Geodesy and Environmental Engineering, AGH University of Science and Technology, Krakow

In spite of these doubtless advantages, splines also suffer from some disadvantages which, in certain cases, considerably hinder their application. The first of these involves the necessity of performing relatively uniformly arranged observations. Otherwise, large errors of approximations are obtained. Another disadvantage is the consumption of time for the operations performed on the spline already generated (e.g. creating cross sections, determining intersections with other structures/objects, etc.). If the building structure is observed by means of a laser scanner that delivers very dense clouds of points, the time for performing such type of operations, even with the use of contemporary computers, may be expressed in terms of tens of minutes or even longer.

Therefore, in the case of building structures surveyed non-uniformly or represented by numerous sets of points, it is more beneficial to recourse to conventional methods which allow obtaining better results even with regards to accuracy and time. They exist in various forms with regards to the application of the least squares method (LSM); however, in various conditions, ambiguous results of approximations may be obtained with these methods.

This paper is aimed at describing currently used methods of point set approximation by means of ideal mathematical surfaces, as well as at the analysis of their suitability depending on the type of the structure being approximated and the complexity of the method.

## 2. The Application of LSM for Approximating the Surfaces of Engineering Structures

Currently, the use of a laser scanner not only for measuring the structure/object selected, but also for the general survey of the infrastructure found in a specific area has become common practice. In the case of performing such measurements in the areas of industrial plants, where there are lots of smaller inter-connected structures in the form of cuboids, cylinders, spheres, etc., there is usually no need to determine their local deformations, but the inter-location of these components is of great importance. Therefore, the use of such approximation methods that may change the dense clouds of points for uncomplicated ideal objects, which significantly simplifies and speeds up performing further analyses, is worth the effort. Various types of LSM assisted approximations, by means of simple mathematical representations (surfaces, spheres, hyperboloids, etc.), are perfectly suitable for this purpose.

The simplest and most commonly used is the minimization of algebraic distance. If the function describing the surface  $F(x, y, z)$  passes accurately through all the points, this distance is equal to zero. Where any of the points fail to lie on the surface, then  $l_i = F(x_i, y_i, z_i) \neq 0$ , where  $l_i$  is the so-called algebraic distance. It is

natural to aim at matching of the surfaces into the set of points such that these distances are as small as possible. In this case, the LSM model is as follows

$$\sum (l_i)_2 = \sum (F(x_i, y_i, z_i))^2 \rightarrow 0 \quad (1)$$

For example, for second-degree surfaces commonly used

$$\begin{aligned} l_i = F(x_i, y_i, z_i) = & a_{11}x_i^2 + a_{22}y_i^2 + a_{33}z_i^2 + \\ & + 2a_{12}x_i y_i + 2a_{13}x_i z_i + 2a_{23}y_i z_i + 2a_{41}x_i + 2a_{42}y_i + 2a_{43}z_i + a_{44} \end{aligned} \quad (2)$$

However, at this moment, we should consider the gains of such approximation. To say the truth, the algebraic distance can not be transformed into a geometrical distance linking the spatial points with the surface specific to it. In the geometrical sense, there is no knowledge of what the algebraic distance really is. Therefore, the parameters of the surface determined based upon this method fail to hold clear interpretation. Furthermore, there is yet another serious disadvantage to this method, notably where approximation is performed in different coordinate systems, the parameters obtained will also differ (they will not be invariant). Hence the question may be asked regarding the purposefulness of using such modelling methods. Practical examples show that, under certain conditions, it is possible to obtain parameters very close to the model minimizing the geometrical distances, while the absence of maintaining their invariability is not significant. These conditions are related to the methods of covering the structure/object with measuring points. Such cover should encompass a larger part of the mathematical surface according to the equation in which the approximation is made. Therefore, if we face a sphere tank which could not be measured, e.g. at the bottom and on top, the matching should bring about with good results. However, where the entire structure is e.g. the small sector of the sphere, the results of matching into may be unpredictable, particularly where it has been subject to certain deformations. The advantage of LSM with the use of minimization of algebraic distances is the exceptional simplicity and speed of calculations, which, with good cover of structures with points, constitutes sufficient prerequisites for the application of this solution.

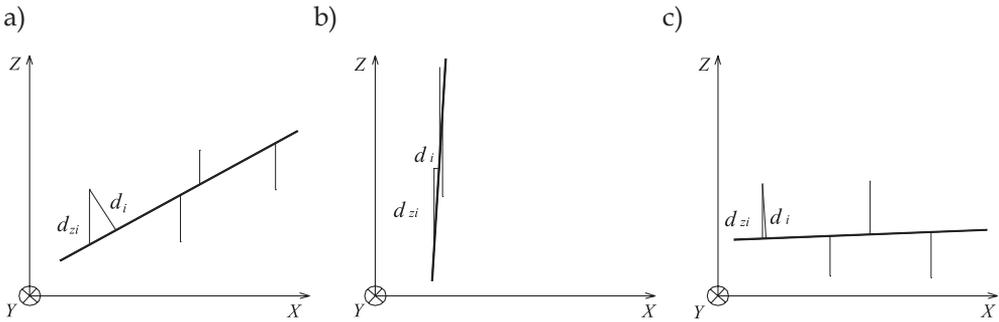
Depending on the type of the surface of the structure and the complexity of calculations, approximations may be performed by means of the functions given in the explicit, implicit or parametric form. The explicit form may be used when approximating flat structures by means of regression planes. The simplest calculation model is then as follows

$$\sum (d_{zi})^2 = \sum (F(x_i, y_i) - f(x_i, y_i))^2 \rightarrow 0 \quad (3)$$

where:

$$\begin{aligned} F(x_i, y_i) = Z_i = ax_i + by_i + c & \quad - \text{ approximating function,} \\ f(x_i, y_i) = z_i & \quad - \text{ the set of values to fit.} \end{aligned}$$

The model clearly shows that in this case, the distances  $d_{zi}$  between the points and the surface (Fig. 1a – plane stretches along Y axis), rather than spatial distances  $d_i$ , are subject of minimization. With the arbitrary orientation of the structure measured in the space, this is unbeneficial as such approximation does not mean the best matching into the set of points. A worst situation will take place where the points are located, as shown in figure 1b. In this case, the distances  $d_{zi}$  have absolutely nothing in common with the lowest distances  $d_i$  from the plane that is best matched into the specific points. What is worse, even a small measuring error or deformation of such surface will bring about very large changes to the minimized  $d_{zi}$ , which, in turn, transforms into totally unpredictable orientation in the space of the plane determined. A simple way out is to make the rotation of the cloud of the measuring points such that it is brought to the approximate parallelism to the plane XY of the coordinate system (Fig. 1c).

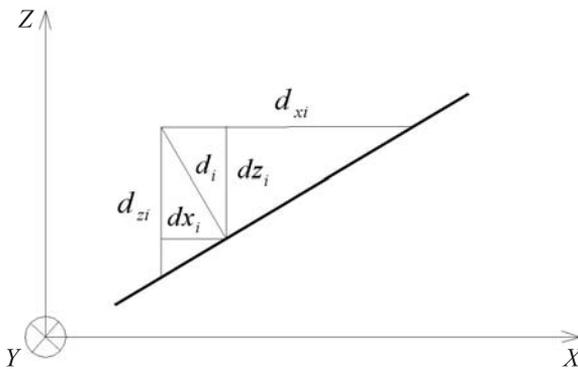


**Fig. 1.** LSM approximation regard of chosen variable  
 Explanations in text

The distances  $d_{zi}$  are then very close to the spatial distances  $d_i$ . LSM assisted approximation performed with their use will always give the results in conformity to approximation along the spatial distances, and the discrepancies, if any, generate changes found within the limits of the measuring error. Provision for parallelism of the cloud of points with plane XY does not need to be very accurate. It is enough to select three points from the cloud which well represent its spatial location (arranged far away from one another, at the best close to the extremes, provided that they fail to display lateral deviations from the cloud) and then to run the interpolating surface through them and determine its inclines with relation to the axes of the coordinate system and perform relevant rotations. There is, however, no need to perform rotations for the points arranged as shown in figure 1b. The fact may be taken advantage of that in this case, the distances  $d_i$  are close to  $d_{xi}$ , and the approximation may be made with regards to axis X rather than Z as before.

This method may also be applied for nonlinear surfaces provided, however, that their curvature is very small (i.e. they may be located in space such that  $d_{zi} \approx d_i$ ). Obviously, the function used for approximation purposes must, in such case, be represented in an explicit form. Attention should be drawn to the fact that parameters to determined are not invariant towards the coordinate system as this is not the spatial distance that is subject to minimization, but the distance along one of the axes, which changes depending on the coordinate system.

Frequently, attempts may be encountered for remedying disturbances related to the LSM for explicit functions in the form of performing approximations with regards to each variable  $x, y, z$  (minimization of  $d_{xi}, d_{yi}, d_{zi}$ ) separately and then determining the conclusive parameters as the average, weighted average, etc. In such a situation, relevant rotations of the clouds of points are not performed. Usually, this method brings some improvement in cases where the cloud of points shows an inclination with regards to all axes of the system; however, for the situation shown in figure 1c, it may only deteriorate the approximation. In this situation, one correct approximation is made along the distances  $d_{zi}$ , comparable to the spatial distances  $d_i$ , and two approximations minimizing the distances  $d_{xi}, d_{zi}$  are made on the rules, as in figure 1b. The conclusive parameters are then determined on the grounds of the approximation giving good results and two approximations that may bring about significant ambiguity. They are not invariant towards the coordinate system either. Yet another factor discouraging the use of this method is the fact that these are not the components  $dx_i, dy_i, dz_i$  of the spatial vector  $d_i$  which are subject to minimization, but the distances  $d_{xi}, d_{yi}, d_{zi}$  along the axes of the coordinate system (Fig. 2) having almost nothing in common with the former.



**Fig. 2.** Components  $dx_i, dy_i, dz_i$  of spatial vector  $d_i$ , distances  $d_{xi}, d_{yi}, d_{zi}$  of point from object  
 Explanations in text

From the methodological point of view, performing the least square method with the use of minimization of geometrical distance  $d_i$  is the proper approach.

This distance is uniquely understood as the spatial distance between the point and the surface; therefore, approximations with the use of this distance provide results whose geometrical interpretation are unique. In addition, in this method, the parameters determined are invariant, irrespective of the system of coordinates where the points have been located.

The application of LSM, similarly to the cases discussed to date, may be described with the following relation

$$\sum (d_i)^2 \rightarrow 0 \quad \text{or} \quad D^T D \rightarrow 0, \quad \text{where} \quad D = \begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_n \end{bmatrix} \quad (4)$$

whose minimization leads to the set

$$2J^T D = 0, \quad \text{where} \quad J = \frac{\partial D}{\partial A} - \text{Jacobian matrix} \quad \partial A = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_k \end{bmatrix} \quad (5)$$

The set of equations allows determining of the  $k$  parameters  $a_i$  of the function approximating the actual shape, given in the point form. Even in the simplest case of matching the straight-line into the set of points, the spatial distance  $d_i$  is non-linear with regards to the parameters determined, and this is why the set (5) is also non-linear. The basic obstruction for applying this method is therefore the complexity and the outlays of calculations, specific to solving non-linear sets of equations. There are many numeric methods allowing to determine the sought parameters  $a_i$  by means of iteration; however, they significantly differ from one another with the speed of convergence to such vector of  $A$  parameters where the equation (5) is fulfilled. Due to rounding errors, it is impossible to determine  $A$  accurately, i.e. by having a sequence of subsequent solutions, we may see that, from a certain moment, it becomes less and less convergent to conclusively stop being diminished before the accurate fulfilment of (5) has been achieved. Usually, however, the requirements regarding the results obtained are not high enough to justify performing iterations for as long as maximum limiting accuracy has been achieved. Speed is not the only criterion of applying these methods, as the width of convergence interval to  $A$  (related to the accuracy of determining the approximate initial solution) is also very important. Depending on this width, particular methods may find inappropriate local extrema of the functions or may become divergent. This is why it is important that the initial vector of solutions be determined as accurately as possible. This will contribute not only to achieving the proper solution, but also it will

be arrived at by means of a smaller number of iterations. To this end, by means of (1), the set of  $k$  interpolation solutions may be solved, determining the vector of  $k$  initial parameters. Usually, methods with a higher rate of convergence have a narrower convergence interval; therefore, combinations of various methods may be applied. For example, initial parameters may be determined by means of (1), their more accurate approximation by means of a method with a wide convergence interval and the conclusive solution by means of a method with a higher rate of convergence. Disturbances to particular methods may involve the necessity of determining e.g. the matrices of second derivative  $d_i$  with respect to parameters  $a_i$  (Hessian matrix), etc. In the case of matching a number of surfaces into the dense set of points originating from laser scanning, the time necessary for solving the entire task needs to be taken into consideration, remembering that some higher rate convergent methods may contain many steps under particular iterations. This becomes particularly significant where the automatic elimination of major errors (M-estimations) is introduced into calculations, which subjects the entire process of determining the parameters  $a_i$  to subsequent iterations removing these errors. It should be remembered that the properties of different methods depend not only on their structure, but also on the form of approximating functions.

Below are mentioned some methods allowing for solving the non-linear set of equations [1, 3, 5, 6], i.e. the implementation of LSM for minimization of the geometrical distance  $d_i$ .

The Newton's method (tangents) is commonly used. Its iteration structure has been presented below

$$(J^T J + HD)\Delta A = -J^T D \quad (6)$$

where:

$$\Delta A = A_i - A_{i-1},$$

$$H = \frac{\partial^2 D}{\partial A^2} \text{ - Hessian matrix.}$$

Rate of convergence of this method is very high (square); however, a disadvantage to this is the necessity of determining the matrices of the second derivatives  $H$ , which increases the cost of the single iteration. In addition, the convergence interval is relatively narrow, which requires the delivery of good, approximate initial parameters.

A modification of the Newton's method, ignoring determination of the Hessian matrix, is the Gauss-Newton method

$$(J^T J)\Delta A = -J^T D \quad (7)$$

Such an assumption may be made only where the value of the relation (4) is close to zero, and due to the fact that aiming at this is the essence of LSM, this method is applied strictly for solving nonlinear tasks of LSM. The advantage to this is the lower cost of performing particular iterations with a similar rate of convergence.

The so-called secant method is encountered even more rarely, and it is used for the troublesome calculations of the Jacobian matrix. However, it is characteristic of a lower rate of convergence than the cases described above, which, as the number of equations and unknowns grows, reduces even further.

The gradient descent method should also be mentioned, as its rate of convergence is low (linear), but its convergence interval is wide, and this is why it is frequently used for determining the initial values of parameters  $a_i$  for Newton's methods.

The Levenberg-Marquardt method, constituting a combination of Gauss-Newton and the gradient descent method, is used most frequently. It is also described by the iteration equation

$$(J^T J - \alpha I) \Delta A = -J^T D \quad (8)$$

where:

- $\alpha$  - parameter of the method selection,
- $I$  - unitary matrix.

For  $\alpha$  approaching zero, the performance of this method approaches the Gauss-Newton method, and for  $\alpha$  increasing, it becomes similar to the gradient descent method. By controlling  $\alpha$  in subsequent iterations, we may safely determine the proper inception values of the vector of the sought parameters  $a_i$  and then, using methods with higher rates of convergence, obtain the conclusive result.

### 3. Summary

The variety of the above-mentioned methods shows how complex the use of LSM is for approximation of the surfaces of engineering structures. Depending on the type of approximating function and the expected final results of matching, various methods may be selected, remembering that the ease and speed of performing calculations are also important. Taking into account the cost-effectiveness aspects, there may be no purpose in using the Levenberg-Marquardt method combined with M-estimation for the purposes of relatively accurate determination of the spatial location of the flat wall of a building. On the other hand, for accuracy aspects, approximation of the sector of rotary hyperboloid should not be performed with

the use of LSM for algebraic distances The main pressure with regard to the proper implementation of these methods is currently put on the creation of software allowing for the matching of various and different objects and structures into the dense clouds of points originating from laser scanning. Currently, full automation of this process is not possible as interaction on the part of the user is required, deciding in more complex cases on the selection of the approximating surface or the methods of determining thereof.

## References

- [1] Cheney W., Kincaid D.: *Analiza numeryczna*. WNT, Warszawa 2006.
- [2] Farin G.: *Curves and surfaces for computer aided deometric design*. Academic Press Inc, New York 1993.
- [3] Fortuna Z., Macukow B., Wąsowski J.: *Metody numeryczne*. WNT, Warszawa 2001.
- [4] Kiciak P.: *Podstawy modelowania krzywych i powierzchni*. WNT, Warszawa 2000.
- [5] Kosma Z.: *Metody numeryczne dla zastosowań inżynierskich*. Politechnika Radomska, Radom 1999.
- [6] Sung J.A.: *Least squares orthogonal distance fitting of curves and surfaces in space*. Springer-Verlag, Berlin 2004.