Optimal Fitting of Curves in Monitoring Data Using the ℓ_1, ℓ_2 and ℓ_∞ Norms

K.E. PARSOPOULOS, V.A. KONTOGIANNI, S.I. PYTHAROULI, P.A. PSIMOULIS, S.C. STIROS and M.N. VRAHATIS, Greece

Key words: curve fitting, ℓ_1 norm, ℓ_2 norm, ℓ_∞ norm, Particle Swarm Optimization, landslide

SUMMARY

Curve fitting is usually based on the least squares method (ℓ_2 norm) and empirical selection of curves using the trial-and error-technique. This approach is simple and successful in the cases measurements contain only random, especially normally distributed errors. In the case of series of monitoring measurements in engineering surveys, to some degree noisy, usually unevenly distributed versus time and including step-like offsets, this approach may lead to very complicated functions, occasionally lacking any clear physical significance and potential for further mathematical analysis. In such cases, however, optimal fitting may alternatively be obtained using simpler functions and the ℓ_1 and ℓ_∞ norms. Interestingly, the use of these norms in the geodetic literature is at least extremely limited, although present-day computers have minimized any limitations in their application.

As an example, fitting of various exponential curves in a geodetic record for monitoring a landslide indicates that the use of the ℓ_1 and ℓ_∞ norms can provide acceptable fitting (i.e. small mean absolute and standard errors) with simpler curves which have a more clear physical significance and higher potential for further analysis than by more complicated curves obtained with conventional least square (ℓ_2 norm) techniques.

Optimal Fitting of Curves in Monitoring Data Using the ℓ_1, ℓ_2 and ℓ_∞ Norms

K.E. PARSOPOULOS, V.A. KONTOGIANNI, S.I. PYTHAROULI, P.A. PSIMOULIS, S.C. STIROS and M.N. VRAHATIS, Greece

1. INTRODUCTION

Geodetic monitoring data usually produce a set of observations in the form of time series, or alternatively of a record of dislocations of control points along a distance (Fig. 1).

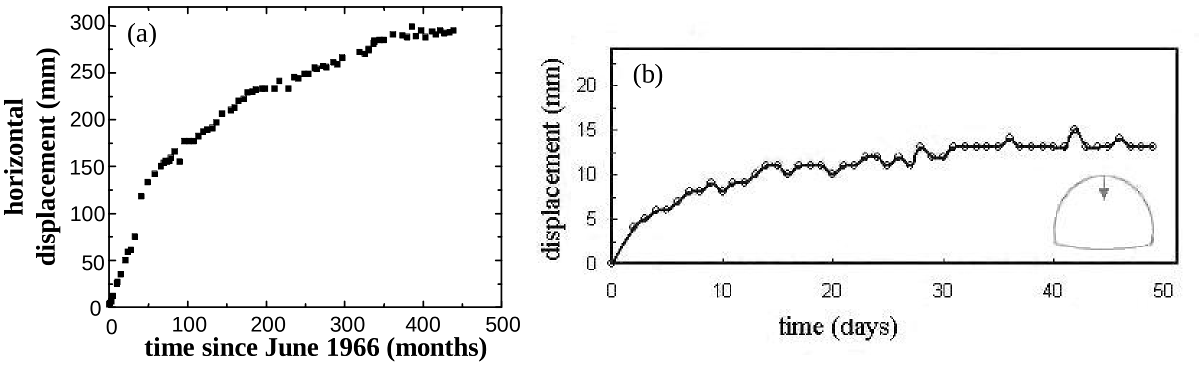

Fig. 1 a: displacement history of a control station at the crest of the Kremasta Dam, Greece. (after Stiros et al. (2004), simplified), **Fig. 1 b:** cumulative displacement along the axis of a control station during the excavation of the Patras by-pass tunnel, Greece (after Kontogianni and Stiros (2004), simplified).

The usual procedure is to approximate such series of measurement (in the following regarded for simplicity as time series) by curves in order (1) to decompose observations into different signals (Akansu and Haddad (2001); Proakis and Manolakis (1996); Pytharouli et al. this volume), or (2) simply to approximate measurements by easy-to-handle mathematical formulas with a clear physical significance (for instance, convergence in a tunnel section tending to stabilization after a gradually decreasing rate of deformation can be approximated by a simple exponential function).

Traditionally, this curve fitting procedure is based on the Least Squares Method (LSQR) which provides unbiased and minimum variance estimates on the condition that measurement errors are random and hopefully normally distributed (Mikhail, 1976; Marshall and Bettel, 1996; Amiri-Simkooei, 2003). This signifies that the error properties of measurements are known, or at least that outliers can be detected and eliminated before the application of the LSQR method.

In the case of geodetic monitoring measurements, this is certainly not an easy task. Field measurements (1) are usually contaminated by various systematic errors, (2) usually reflect the superposition of various signals, (3) are in many cases unevenly distributed versus time (or distance, etc.), and (4) contain either isolated outliers which can rather easily be identified using a rejection limit (usually the 3sigma criterion) or more elaborate techniques (Marshall and Bethel, 1996; Amiri-Simkooei, 2003), or a number of measurements with values quasi-normally distributed between mean values and outliers; the significance of such measurements is not clear, and the way to be treated is therefore rather subjective. For these reasons, the calculated error estimates of certain time series seem not realistic, and empirical estimates, occasionally 15 times the computed ones are sometimes adopted by careful investigators (Cocard et al., 1999). Obviously, in such cases, analysis with LSQR does not have the typical properties of this method.

Another aspect of the problem of curve fitting in monitoring data is that investigators are confined to the empirical selection of types of curves to be fitted on the basis of the minimum variance criterion and the LSQR method, i.e. to one degree of freedom in their analyses. Because of the error properties of monitoring data mentioned above, usually reflected in steps and offsets, especially frequent in geotechnical and geologic studies, this approach necessarily leads to rather complicated types of curves (i.e. to a combination of polynomials of high degrees, exponential curves, etc.), otherwise the fit is not good. Obviously, such curves cannot be easily analyzed and processed; for instance, if it can be shown that the general trend in a certain phenomenon, for instance a landslide or the convergence in a tunnel can be described by a simple exponential curve (Fig. 1), the ratios of curves of different points could perhaps be easily correlated with some other variables defining geotechnical properties of the rocks and obtain a certain analytical relationship between kinematic and geotechnical characteristics of the rocks, structures, etc; this is definitely not an easy task if, instead of a simple exponential curve, a more complicated curve is to be analyzed.

This discussion leads to the question: given the error properties of many geodetic time series, except for the LSQR, is there an alternative method that may permit an optimal curve fitting, i.e. a two-degrees of freedom in the curve fitting problem? In the following, we shall show that there exist alternative approaches, quite realistic with the present-time computers; in particular, the ℓ_1 and ℓ_∞ norms; note that LSQR is recognized as the ℓ_2 norm.

2. THE ℓ_1, ℓ_2 And ℓ_∞ NORMS

The basic design of a curve fitting is to minimize residuals ε_ι, i.e. differences between observed and estimated (predicted) values in all measurement points.
This is a problem with infinite solutions, and among them the ℓ_2 norm solution, known as least square (LSQR), is traditionally the most popular. A main reason for that is that it is based on formulas that can be easily differentiated, in contrast with all alternative approaches which are based on computer approximation techniques. Among these last approaches, ℓ_1, and ℓ_∞ norm solutions are the most common ones (Liu and Chen, 1998).

The mathematical definition of these solutions is given below, assuming a variance-covariance matrix equal to the identity matrix, an assumption reasonable for data collected under standard conditions.

$$\|\varepsilon\|_1 = \sum_{i=1}^{m} |\varepsilon_i|, \qquad \|\varepsilon\|_2 = \left(\sum_{i=1}^{m} |\varepsilon_i|^2 \right)^{1/2}, \qquad \|\varepsilon\|_\infty = \max_{1 \le i \le m} |\varepsilon_i|$$

where m is the number of observations and $\varepsilon_i = M_i - O_i$, i = 1, …,m with O_i being the i-th observed value and M_i be the corresponding value implied by the model.

Obviously, norm ℓ_1 *is a priori* the most reasonable choice, for the basic aim in curve fitting is to minimize residuals; this solution provides unbiased results, but no minimum variances (Marshall and Bettel, 1996; Amiri-Simkooei, 2003).

On the other hand, ℓ_∞ norm has the advantage of optimal approximation of values deviating from the average trend, and hence its results are superior in case for instance of edge points of a segmented time-series.

Various approximation techniques have been proposed for curve fitting based on ℓ_1 and ℓ_∞ norms (see for instance Amiri-Simkooei, 2003). In our computations we have, however, adopted an approach that incorporates the Particle Swarm Optimization algorithm (Eberhart and Kennedy, 2001), which has been applied on several curve fitting problems with success (Parsopoulos et al., 2001).

3. A CASE STUDY

The one and two-degrees of freedom approach in curve fitting is examined in the case of a long monitoring record of a landslide. The distance of a selected point close to the crown of the landslide from a reference point on stable ground was systematically measured using EDM and a uniform measuring procedure for 20 years. The accuracy of this record, containing long and shorter periods of no measurements and consisting of 480 surveys in total (Fig. 2), was evaluated on the basis of high accuracy leveling and was found to contain some outliers which were discarded. The remaining, acceptable values seem not to correspond to a rather smooth curve, but are characterized by step-like offsets, indicative of periods of acceleration of the movement of the landslide as a consequence of seismic, hydrological and other effects.

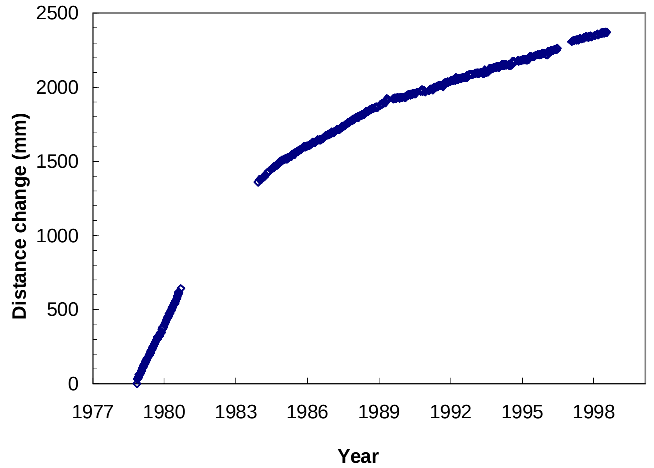

Fig. 2: Time series of the distance change of a control station of a landslide in Greece from a reference station (after Stiros et al., 2004)

This record was modeled by two different types of exponential curves,

Model 1: $f(t) = A(1 - e^{-\frac{t}{B}}) + C$

Model 2: $f(t) = A(1 - e^{-\frac{t}{B}}) + Kt + C$

and three different approximation methods based on the ℓ_1, ℓ_2 and ℓ_∞ norms.
The results of these approximations are shown in Fig. 4 and Table 1.

It is evident that the fitting of the mathematically more complicated model 2 is better than model 1 by a factor of 2. However, the linear terms in this model make any comparative analysis of the displacement of the various control points, not examined here, and of other parameters a difficult task. This situation would deteriorate in the case of additional terms which would typically refine fitting.

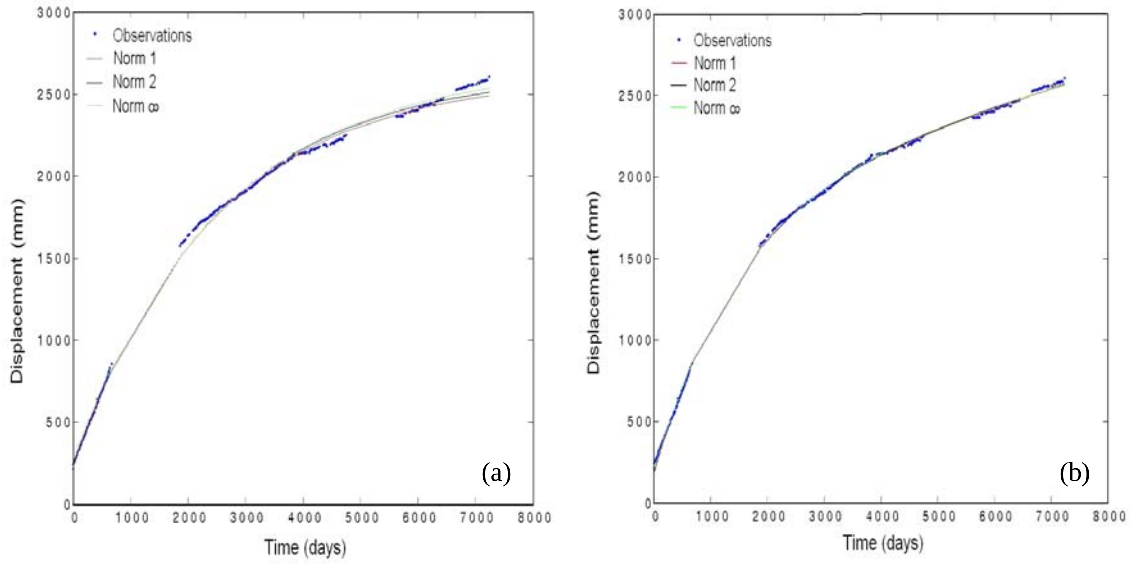

Fig. 4: Results of fitting using three different approximation methods based on the ℓ_1, ℓ_2 and ℓ_∞ norms for a: Model 1, b: Model 2

Concerning model 1, fitting based on the ℓ_1 norm is better than that obtained through LSQR (norm ℓ_2), on the basis of the mean absolute error, while on the basis of the standard error, fitting on the basis of the ℓ_∞ norm is better (Fig.5a). Apart from that, this last fitting seems superior as far as the edge points of the segments are concerned (Fig. 4b).

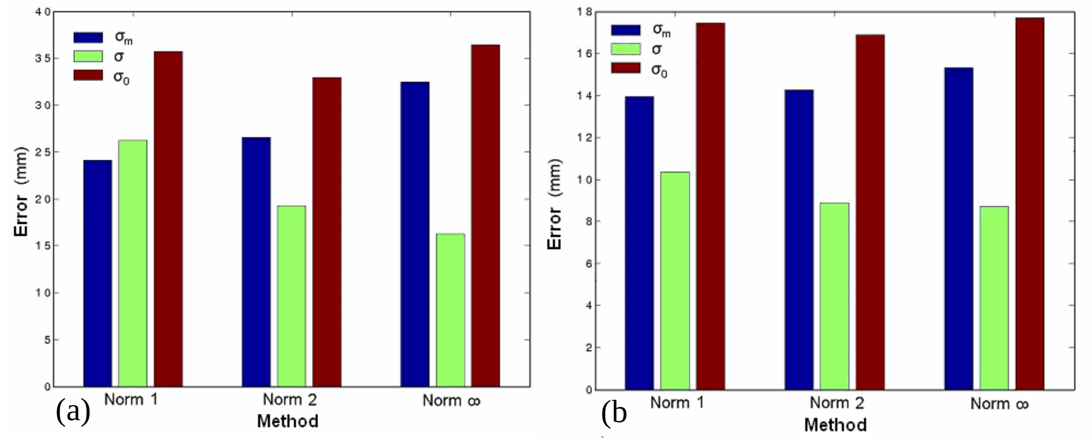

Fig. 5: Mean absolute error σ_m, standard deviation of the absolute error ς and standard error of each observation after the fitting of a: model 1 and b: model 2 to the observations using the ℓ_1, ℓ_2 and ℓ_∞ norm.

Table 1: Estimated values of coefficients of Model 1 and Model 2

Model	Norm	A	B	C	K
Model 1	ℓ_1	2377.10	2447.27	236.68	-
	ℓ_2	2391.74	2468.11	246.08	-
	ℓ_∞	2397.33	2588.62	283.42	-
Model 2	ℓ_1	1626.50	1431.98	0.106	178.76
	ℓ_2	1624.57	1464.39	0.106	187.87
	ℓ_∞	1686.59	1582.49	0.096	202.47

4. CONCLUSIONS

Recent advances in computers and arithmetic analyses give the opportunity of two degrees of freedom in curve fitting in time series (selection of the curve model and of the approximation technique), instead of the one degree of freedom (selection of a curve model only) based on LSQR.

Optimal curve fitting is therefore possible for curves corresponding to simpler and easier to analyze mathematical models. This gives the opportunity of further analysis of monitoring data and their correlation with geotechnical, geological, etc data and perhaps the identification of new or refinement of relationships between various physical parameters.

REFERENCES

Akansu, A.N. and R.A. Haddad (2001), "Multiresolution Signal Decomposition: Transforms, Subbands, Wavelets", Academic Press, CA

Amiri-Simkooei A. (2003), "Formulation of L_1 Norm Minimization in Gauss-Markov Models", *Journal of Surveying Eng.*, 129, 1, pp. 37-43

Cocard, M, H.-G. Kahle, Peter, Y., Veis, G., Felekis, S., Paradissis, D. and H. Billiris (1999), "New constraints on the rapid crustal motion of the Aegean region: recent results inferred from GPS measurements (1993-1998) across the Hellenic Arc, Greece", *Earth and Planetary Science Letters*, 172 (1-2), pp. 39-47.

Kennedy, J. and R.C. Eberhart (2001), "Swarm Intelligence", Morgan Kaufmann Publishers

Kontogianni, V. and S. Stiros (2002), "Shallow tunnel convergence: Predictions and Observations". *Engineering Geology*, 63 (3-4), pp. 333-345.

Kontogianni, V., Tzortzis, A. and Stiros, S. (2004). Deformation and Failure of the Tymfristos Tunnel, Greece. *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE (in press).

Liu, Q.W. and Y.Q. Chen (1998) "Combining the Geodetic Models of Vertical Crustal Deformation", *Journal of Geodesy*, 72, pp. 673-683

Marshall, J. and J. Bethel, (1996), "Basic Concepts of L_1 Norm Minimization for Surveying Applications", *Journal of Surveying Eng.*, 122, 4, pp. 168-179

Mikhail, E. (1976), "Observations and least squares", IEP, New York, NY

Parsopoulos, K.E., E.C. Laskari and M.N. Vrahatis (2001), "Solving l_1 norm errors-in-variables problems using particle swarm optimizer", In M.H. Hamza, editor, Artificial Intelligence and Applications, pp. 185-190, IASTED/ACTA Press

Proakis, J.G and D.G. Manolakis (1996), "Digital Signal Processing", Prentice-Hall, NJ

Pytharouli, S., V. Kontogianni, P. Psimoulis and S. Stiros (2004), "Spectral Analysis Techniques in Deformation Analysis Studies", this volume

BIOGRAPHICAL NOTES

Konstantinos E. Parsopoulos, Dipl. Math., MSc, is PhD candidate of the Dept. of Mathematics, University of Patras, Greece. His research interests include among others development and analysis of evolutionary and swarm intelligence algorithms for optimization and applications.

Villy A. Kontogianni, Dipl. Eng. MSc, is PhD candidate of the Dept. of Civil Engineering, Patras University, Greece. Her PhD thesis is the analysis of tunnel deformations monitoring data. Her research, in the Geodesy Lab., involves analysis of geodetic data from geotechnical engineering structures (tunnels, dams, ground subsidence) and seismic faults activation.

Stella I. Pytharouli, Dipl. Eng. MSc, is a PhD student of the Dept. of Civil Engineering, Patras University, Greece. Her research activities in the Geodesy Lab., involve analysis of geodetic data from geotechnical engineering structures with main interest on dam deformation and GPS monitoring data analysis.

Panos A. Psimoulis, Dipl. Eng., is a postgraduate student of the Dept. of Civil Engineering, Patras University, Greece. His research activities, involve the application of analytical methodologies on reconstruction/ design of structures and application of GPS in monitoring civil engineering structures.

Stathis C. Stiros, Dipl. Eng. PhD, is Ass. Prof. and director of the Geodesy Lab., Dept. of Civil Engineering, Patras University. His research interests include among others deformation monitoring and analysis surveys in the field of civil engineering and geology/geophysics.

Michael N. Vrahatis, is Prof. with the Dept. of Mathematics, University of Patras. His research interests include among others development and analysis of Computational Intelligence algorithms and applications.

CONTACTS

Konstantinos E. Parsopoulos
Dept. of Mathematics,
University of Patras Artificial Intelligence Research Center (UPAIRC)
University Campus
26500 Patras
GREECE
Tel. + 30 2610 997348
Fax + 30 2610 992965
Email: kostas@math.upatras.gr

Ms Villy A. Kontogianni
University of Patras
Geodesy Lab., Dept. of Civil Engineering
University Campus
26500 Patras
GREECE
Tel. + 30 2610 996511
Fax + 30 2610 997877
Email: vkont@civil.upatras.gr

Ms Stella I. Pytharouli
University of Patras
Geodesy Lab., Dept. of Civil Engineering, University Campus
26500 Patras
GREECE
Tel. + 30 2610 996511
Fax + 30 2610 997877
Email: spitha@upatras.gr

Panos A. Psimoulis
University of Patras
Geodesy Lab., Dept. of Civil Engineering, University Campus
26500 Patras
GREECE
Tel. + 30 2610 996511
Fax + 30 2610 997877
Email: papsimouli@upnet.gr

Stathis C. Stiros
University of Patras
Geodesy Lab., Dept. of Civil Engineering, University Campus
26500 Patras
GREECE
Tel. + 30 2610 996511
Fax + 30 2610 997877
Email: stiros@upatras.gr

Michalis N. Vrahatis
Dept. of Mathematics
University of Patras Artificial Intelligence Research Center (UPAIRC)
University Campus
26500 Patras
GREECE
Tel. + 30 2610 997374
Fax + 30 2610 992965
Email: vrahatis@math.upatras.gr