Soft Computing Ideas Can Help Earthquake Geophysics

Solymar Ayala, Aaron Velasco
Department of Geological Sciences
University of Texas at El Paso
El Paso, Texas 79968, USA
sayalacortez@miners.utep.edu, aavelasco@utep.edu

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, Texas 79968, USA
vladik@utep.edu

Abstract—**Earthquakes can be devastating, thus it is important to gain a good understanding of the corresponding geophysical processing. One of the challenges in geophysics is that we cannot directly measure the corresponding deep-earth quantities, we have to rely on expert knowledge, knowledge which often comes in terms of imprecise ("fuzzy") words from natural language. To formalize this knowledge, it is reasonable to use techniques that were specifically designed for such a formalization – namely, fuzzy techniques, In this paper, we formulate the problem of optimally representing such knowledge. By solving the corresponding optimization problem, we conclude that the optimal representation involves using piecewise-constant functions. For geophysics applications, this means that we need to go beyond tectonic plates to explicitly consider parts of the plates that move during the earthquake. We argue that such an analysis will lead to a better understanding of earthquake-related geophysics.**

I. Specifics of Data Processing in Earthquake Analysis (and in Geophysics in General)

Earthquake analysis is important. Earthquakes can be devastating. It is therefore important to gain as much understanding about the corresponding geophysical processes as possible; see, e.g., [1], [7].

Usual approach to data processing. A good understanding means that we know, for each location (x, y) and for each depth z, what is the density ρ at this location and this depth, what are the mechanical properties of the material at this 3-D location $u = (x, y, z)$, what are the stresses at this 3-D location. In other words, we need to find the corresponding functions like

$$\rho(u) = \rho(x, y, z).$$

How can we describe a function in a computer? A usual way is to select a *basis* of functions

$$e_1(u), e_2(u), \ldots, \tag{1}$$

so that each desired function $f(u)$ can be represented as a linear combination of the basis functions

$$f(u) = c_1 \cdot e_1(u) + c_2 \cdot u_2(u) + \ldots,$$

and then represent the desired function $f(x)$ by the corresponding coefficients

$$(c_1, c_2, \ldots). \tag{2}$$

In principle, we can consider different bases, but it is usually convenient to *orthonormalize* them, i.e., to consider linear combinations

$$e_i^{on}(x) = \sum_{j=1}^{i} c_{ij} \cdot e_j(u)$$

for which, for all i and j, we have

$$\int \left(e_i^{on}(u) \right)^2 \, du = 1$$

and

$$\int e_i^{on}(u) \cdot e_j^{on}(u) \, du = 0 \text{ when } i \neq j.$$

In this case, the desired coefficients c_i can be obtained by using a simple formula

$$c_i = \int f(u) \cdot e_i^{on}(u) \, du. \tag{3}$$

Thus, without losing generality, we can safely assume that the basis (1) is orthonormal.

The most widely used examples of such bases are:
- sines and cosines, and
- wavelets; see, e.g., [2], [4], [6], [11].

For sines and cosines, the expansion into the corresponding basis is known as *Fourier transform*. For wavelets, wavelet, the transformation from the original function $f(u)$ to the coefficients c_i is known as the *wavelet transform*.

It is important to select an appropriate basis. It is known that selecting an appropriate basis can drastically improve the quality of the data processing results.

For example, in many cases, wavelet analysis has led to interesting discoveries that were not possible when Fourier analysis was used to process the corresponding data.

It is therefore very important, in each practical situations, to select the most appropriate basis.

What we plan to do in this paper: main idea. In this paper, we provide arguments for selecting the most appropriate basis for earthquake-related analysis. In this analysis, we use the specific features of the geophysical data processing.

Specifics of geophysical data processing. In comparison with most other data processing situations, geophysical analysis has two important specifics.

First, in most data processing situations, we have continuous functions. For example, when we control a vehicle, its location continuously depends on time. In contrast, in geophysics, there are clear discontinuities:

- as we go deeper,
- we have an abrupt transition between different layers.

The second difference is that in most other data processing situations, we can determine the ground truth, i.e., the actual values of the corresponding quantities. In geophysics, our ability to get the ground truth is very limited: up to a certain depth, we can drill a borehole and find out what are the actual properties, but at larger depths, this is not practically possible.

Why soft computing. Since we cannot determine the actual values to check different models, we have to rely on expert knowledge to decide which model works better.

Expert knowledge rarely comes in precise terms, it usually comes in terms of imprecise ("fuzzy") words. To describe the corresponding knowledge in precise terms, it is therefore reasonable to use techniques specifically designed to handle such knowledge – namely, the techniques of fuzzy logic; see, e.g., [3], [5], [8], [9], [10], [12], [13].

II. ANALYSIS OF THE PROBLEM

Main idea. Since the values $f(u)$ comes from expert estimates, they come with a fuzzy uncertainty. In other words, for every u, we have fuzzy information about the difference $\Delta f(u) \stackrel{\text{def}}{=} \widetilde{f}(u) - f(u)$ between:

- the expert estimate $\widetilde{f}(u)$ and
- the actual (unknown) value $f(u)$.

In precise terms, this means that:

- we do not know the probabilities of different possible values of $\Delta f(u)$, but
- we have a membership function $\mu(\Delta f)$ that describes, for each possible value Δf, the degree to which this value is possible.

Since we have no reason to assume that the estimation errors are positive or negative, it is reasonable to assume that the degree of possibility of each value Δf does not depend on its sign:

$$\mu(-\Delta f) = \mu(\Delta f).$$

Once we have selected the basis $e_i(u)$, we will then transform the estimate for $f(u)$ into the sequence of the corresponding coefficients c_i.

- Since the values $f(u)$ are known with uncertainty,
- as a result, we can only determine the coefficients c_i – and thus, the corresponding term $c_i \cdot e_i(u)$ – with uncertainty.

A reasonable idea is to select the basis $e_i(u)$ for which this uncertainty in the term $c_i \cdot e_i(u)$ is the smallest possible.

Let us describe this idea in precise terms. When we process the expert estimates $\widetilde{f}(u)$, we get the following estimates $\widetilde{c}_i$ for the coefficients c_i:

$$\widetilde{c}_i = \int \widetilde{f}(u) \cdot e_i(u) \, du. \tag{4}$$

The actual (unknown) value c_i of the corresponding coefficient can be obtained if we apply the same procedure to the actual (unknown) function $f(u)$:

$$c_i = \int f(u) \cdot e_i(u) \, du. \tag{5}$$

If we subtract (5) from (4) and take into account that the integral of the difference is equal to the difference of the integrals, we get the following formula for the inaccuracy $\Delta c_i \stackrel{\text{def}}{=} \widetilde{c}_i - c_i$:

$$\Delta c_i = \int \Delta f(u) \cdot e_i(u) \, du. \tag{6}$$

The inaccuracy in the product $c_i \cdot e_i(u)$ is equal to the product $\Delta c_i \cdot e_i(u)$. This value depends on the location u:

- for some locations u, the value $|e_i(u)|$ is larger, so the inaccuracy is larger;
- for other locations u, the value $|e_i(u)|$ is smaller, so the inaccuracy is smaller.

It is reasonable to minimize the worst-case inaccuracy

$$\Delta c_i \cdot \max_u |e_i(u)| = \max_u |e_i(u)| \cdot \int \Delta f(u) \cdot e_i(u) \, du. \tag{7}$$

Here, each value $\Delta f(u)$ is a fuzzy number, so Δc_i is also a fuzzy number. In fuzzy logic, this fuzzy number is determined by the *Zadeh's extension principle*.

It is known that in general, computing the result $Y = f(X_1, \ldots, X_n)$ of applying a function $f(x_1, \ldots, x_n)$ to n fuzzy numbers $X_1, \ldots, X_n$ can be described as follows: for each $\alpha \in (0, 1]$, the α-cut

$$Y(\alpha) \stackrel{\text{def}}{=} \{y : \mu(y) \geq \alpha\}$$

is equal to the range of the function $f(x_1, \ldots, x_n)$ when each x_i takes values from the corresponding α-cut

$$X_i(\alpha) \stackrel{\text{def}}{=} \{x_i : \mu_i(x_i) \geq \alpha\}.$$

In precise terms, we have

$$Y(\alpha) = \{f(x_1, \ldots, x_n) : x_1 \in X_1(\alpha), \ldots, x_n \in X_n(\alpha)\}.$$

Let us apply this general result to our formula (7). Since the membership function $\mu(\Delta f)$ does not change if we change the sign of the difference Δf, for each α, the corresponding α-cut is a symmetric interval. Let us denote this interval by

$$[-\Delta(\alpha), \Delta(\alpha)].$$

The expression (7) is a linear combination of all the values $\Delta f(u)$:

- when $e_i(u) > 0$, this function is increasing in $\Delta f(u)$;
- when $e_i(u) < 0$, this function is decreasing in $\Delta f(u)$.

Thus, when each value $\Delta f(u)$ takes all possible values from the interval $[-\Delta(\alpha), \Delta(\alpha)]$, the largest possible value of the expression (6) is attained when:

- for those u for which $e_i(u) > 0$, the value $\Delta f(u)$ is the largest possible, i.e., $\Delta f(u) = \Delta(\alpha)$, and

- for those u for which $e_i(u) < 0$, the value $\Delta f(u)$ is the smallest possible, i.e., $\Delta f(u) = -\Delta(\alpha)$.

In both cases, the largest possible value of the product

$$\Delta f(u) \cdot e_i(u)$$

is equal to $\Delta(\alpha) \cdot |e_i(u)|$. Thus, the largest possible value of the integral (6) is equal to

$$\int \Delta(\alpha) \cdot |e_i(u)| \, du = \Delta(\alpha) \cdot \int |e_i(u)| \, du.$$

Hence, the largest possible value of the integral (7) is equal to

$$\Delta(\alpha) \cdot \max_u |e_i(u)| \cdot \int |e_i(u)| \, du.$$

Similarly, we can show that the smallest possible value of the expression (7) is equal to

$$-\Delta(\alpha) \cdot \max_u |e_i(u)| \cdot \int |e_i(u)| \, du.$$

Thus,

$$Y(\alpha) = [-y(\alpha), y(\alpha)],$$

where

$$y(\alpha) \stackrel{\text{def}}{=} \Delta(\alpha) \cdot \max_u |e_i(u)| \cdot \int |e_i(u)| \, du.$$

The estimate for c_i is the most accurate when this interval is the narrowest possible, i.e., when the value

$$\Delta(\alpha) \cdot \max_u |e_i(u)| \cdot \int |e_i(u)| \, du$$

is the smallest possible.

In this product, the factor $\Delta(\alpha)$ is given. So, the smallest possible value of the above product is attained when the product

$$\max_u |e_i(u)| \cdot \int |e_i(u)| \, du \tag{9}$$

attains its smallest possible value. Hence, we arrive at the following optimization problem.

Resulting optimization problem. Among all possible functions $e_i(u)$ for which $\int e_i^2(u) \, du = 1$, we need to find the function with the smallest possible value of the product (9).

Analysis of the resulting optimization problem. We always have $e_i^2(u) = |e_i(u)|^2$. Thus,

$$\int |e_i(u)|^2 \, du = \int e_i^2(u) \, du = 1.$$

Also, for every u, we have

$$e_i(u) \le \max_v |e_i(v)|.$$

Hence,

$$|e_i(u)|^2 \le \left(\max_v |e_i(v)| \right) \cdot |e_i(u)|. \tag{10}$$

Integrating both parts of this inequality, we conclude that

$$1 = \int |e_i(u)|^2 \, du \le \max_v |e_i(v)| \cdot \int |e_i(u)| \, du. \tag{11}$$

Thus, the product (9) that we want to minimize cannot be smaller than 1. One can easily check that when $|e_i(u)| = \text{const}$ for all u from the given region, we get exact equality in the formula (1) and thus, in formula (11).

So, when the absolute value $|e_i(u)|$ is constant, we attain the smallest possible value 1 of the desired product (9).

Vice versa, if at least for one value u, we have strict inequality in (10), we will have strict inequality in (11) as well. So, to attain the smallest possible value of the product (9), we must always have equality in the formula (1), i.e., we must always have the following equality:

$$|e_i(u)|^2 = \left(\max_v |e_i(v)| \right) \cdot |e_i(u)|. \tag{12}$$

When $e_i(u) \ne o$, we can divide both sides of this equality by $|e_i(u)|$ and conclude that

$$|e_i(u)| = \max_v |e_i(v)|.$$

In other words, for every u:
- we either have $e_i(u) = 0$
- or we have $|e_i(u)|$ equal to the largest possible value

$$m \stackrel{\text{def}}{=} \max_v |e_i(v)|.$$

So, we arrive at the following conclusion:

Resulting solution. Each function $e_i(u)$ from the geophysically optimal basis take only three values: 0, m, and $-m$, for some real number $m > 0$.

This means, in particular, that all the optimal basis functions are piecewise constant.

Comment. Let us consider the geophysical meaning of this result.

III. GEOPHYSICAL MEANING OF OUR RESULT

What does our result means in terms of earthquake analysis. An earthquake leads to a spatial shift at different locations. For catastrophic earthquakes, this shift can be in meters; for smaller earthquakes, we can have centimeters-size shift.

In general, we have a shift $s(x, y)$ as a function of 2-D spatial coordinates. Our optimization result shows that the optimal way to analyze the empirical data about this shift is to represent it as a linear combination $\sum c_i \cdot e_i(u)$ of piece-wise constant functions $e_i(u)$. Such a linear combination is also piece-wise constant. Thus, what we need to do is divide the whole area into several zones, in each of which the shift is fixed.

In geometric terms, this means that instead of considering each spatial location (x, y) by itself, we divide the whole region into parts, each of which moves as a whole (i.e., as a solid body).

How do we transform the observed shifts into this piece-wise constant presentation: an algorithm. When a function is piece-wise constant, it means that it attains finitely many

different values. Let us sort these values into an increasing order:

$$s_1 < s_2 < \ldots < s_m.$$

Suppose at first that these values are given. In this case, we want to approximate the original function $f(u)$ by a piece-wise constant function $a(u)$ that takes values s_i. For each u, the value $a(u)$ is equal to one of the values $s_1, \ldots, s_m$. Thus, describing the function $a(u)$ is equivalent to describing, for each i from 1 to m, the set S_i of all the locations u for which $a(u) = s_i$. These m sets should form a partition of the original domain S, i.e., we should have

$$S_1 \cup \ldots \cup S_m = S$$

and

$$S_i \cap S_j = \emptyset$$

for all $i \neq j$.

A natural idea is to use the Least Squares approach, i.e., to find such a function $a(u)$ for which the integral

$$\int (f(u) - a(u))^2 \, du$$

attains the smallest possible value. One can easily check that the integral attains the smallest possible value if and only if for each u, we select the value $a(u) \in \{s_1, \ldots, s_m\}$ for which the value $(f(u) - a(u))^2$ is the smallest possible. In other words, for each location u, as $a(u)$, the take the value s_i which is the closest to the original value $f(u)$. In other words:

- we select $a(u) = s_1$ if

$$f(u) \leq \frac{s_1 + s_1}{2};$$

- we select $a(u) = s_2$ if

$$\frac{s_1 + s_2}{2} \leq f(u) \leq \frac{s_2 + s3}{2};$$

- ...
- for each $i = 2, \ldots, m-1$, we select $a(u) = s_i$ if

$$\frac{s_{i-1} + s_i}{2} \leq f(u) \leq \frac{s_i + s_{i+1}}{2};$$

- ...
- finally, we select $a(u) = s_m$ if

$$\frac{s_{m-1} + s_m}{2} \leq f(u).$$

We can repeat this procedure for different tuples

$$s = (s_1, \ldots, s_m).$$

For each such tuple, we find the resulting mean square error

$$\int \min_i (f(u) - s_i)^2 \, du.$$

We then select a tuple s for which this mean square error attains the smallest possible value.

This immediately brings to mind tectonic plates. The above piece-wise description bring to mind the geophysical idea that the earth's surface consists of tectonic plates, solid bodies that move in relation to each other.

So, in the first approximation, our mathematical result leads to the very well-known plate tectonics idea.

Our result goes beyond plate tectonics. In the first approximation, our result simply leads to a well-known idea of plate tectonics. In this first approximation, the whole plate moves as a whole, the shift is exactly the same on all locations from this plate.

In practice, the shift is somewhat different in different locations on the same tectonic plate. To capture this difference and thus, provide a more accurate description of the corresponding geophysics, we therefore need to divide each affected plate into two (or more) different parts, with different shifts in each part.

This idea has geophysical sense. It is known that the major earthquakes are caused by the interaction of tectonic plates – that move relative to each other. As a result, all major earthquakes – and the vast majority of smaller earthquakes – happen at the boundaries between tectonic plates. Specifically, they happen at the *convergent* boundaries, where the plates move towards each other, accumulating a stress. This stress is released by an earthquake.

The above description is a first crude approximation to the corresponding geophysics, in which we can consider the whole plate as a solid body, in which all parts move the same way. In reality, different parts of the plate may accumulate the stress differently and move differently. As a result, some earthquakes only involve a part of the boundary between the plates. Depending on the size of this part, we can get earthquakes of different magnitudes.

Beyond piece-wise constant functions: geophysics-motivated idea. Solid bodies do not just shift, they can also rotate. So, a natural idea is to consider not only shifts, but also rotations of the parts of the plate.

In this case:

- instead of approximate the measured values $f(u)$ by a piece-wise constant function,
- we approximate it by a piece-wise linear functions corresponding to shifts and rotations of different parts of each tectonic plate.

This can help in earthquake studies. In view of the above, to get a better understanding of the earthquake geophysics, it is important to analyze which parts of the plate are involved in different earthquakes, which parts have accumulated more stress and in which part, the stress has been released.

This idea is challenging. From the computational viewpoint, our idea is very challenging:

- while we can relatively easily identify the boundary between the plates, where the big motion occurs,

- it is much more challenging to identify the parts of the plate that are involved in an earthquake.

The reason why this identification is not easy is because we are interested in geophysical processes far away from the boundaries, where the earthquake-related motion is much smaller in amplitude and thus, much more difficult to detect.

We all need to work together to overcome these challenges. As of now, what we have is ideas and models.

Our preliminary results show that these ideas are promising, and we will continue working on them.

However, we think that it will be beneficial to publicize these ideas so that others can implement them, use them, improve them if needed – and thus, help to get a better understanding of earthquake-related geophysics.

REFERENCES

[1] K. Aki and P. Richard, *Quantitative Seismology, Theory and Methods*, Universoty Science Books, Sansalito, California, 2009.

[2] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, CRC Press, Boac Raton, Florida, 2017.

[3] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[4] E. Chandrasekhar, V. P. Dimri, and V. M. Gadre (eds.), *Wavelets and Fractals in Earth System Sciences*, CRC Press, Boca Raton, Florida, 2014.

[5] A. Hajan and P. Styles, *Application of Soft Computing and Intelligent Methods in Geophysics*, Springer, Cham, Switerland, 2018.

[6] B. Han, *Framelets and Wavelets: Algorithms, Analysis, and Applications*, Springer, Cham, Switzerland, 2017.

[7] J. Havskov and L. Ottemoller, *Routine Data Processing in Earthquake Seismology: With Sample Data, Exercises and Software*, Springer, Dordrecht, Heidelberg, London, New York, 2010.

[8] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[9] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[10] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[11] P. Nicholas, *Framelets and Wavelets: Algorithms, Analysis, and Applications*, Cambridge University Press, Cambridge, UK, 2017.

[12] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[13] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.