TREATMENT OF GEODETIC SURVEY DATA AS FUZZY VECTORS

G. Navratil [*], E. Heer, J. Hahn

Department for Geodesy and Geoinformation, TU Vienna, Austria – (navratil,hahn)@geoinfo.tuwien.ac.at,
elsa.heer@tuwien.ac.at

Commission II, WG II/4

KEY WORDS: Fuzzy Vector, Geodetic Survey Data, Analysis

ABSTRACT:

Geodetic survey data are typically analysed using the assumption that measurement errors can be modelled as noise. The least squares method models noise with the normal distribution and is based on the assumption that it selects measurements with the highest probability value (Ghilani, 2010, p. 179f). There are environment situations where no clear maximum for a measurement can be detected. This can happen, for example, if surveys take place in foggy conditions causing diffusion of light signals. This presents a problem for automated systems because the standard assumption of the least squares method does not hold. A measurement system trying to return a crisp value will produce an arbitrary value that lies within the area of maximum value. However repeating the measurement is unlikely to create a value following a normal distribution, which happens if measurement errors can be modelled as noise. In this article we describe a laboratory experiment that reproduces conditions similar to a foggy situation and present measurement data gathered from this setup. Furthermore we propose methods based on fuzzy set theory to evaluate the data from our measurement.

1. INTRODUCTION

Methods for geodetic observations are typically designed in a way that the resulting errors are small in relation to the observed value and bias is eliminated. Thus the assumption is that the resulting error (the noise) is normally distributed with an expectation value of zero. This requires not only carefully developed equipment but also controlled environmental conditions. With human observers it is possible to perform the observations only under reasonable conditions and postpone observations otherwise. This is not possible with automated systems. Such systems are typically installed if fast response times are essential, for example in case of landslides. Problems occur if the environmental conditions are bad.

If only some of the automatically determined measurements are affected by poor environmental conditions, then the application of robust estimation may be able to identify observations with poor quality and provide good results by eliminating observations with a bad quality. However, if the observed situation changes during the poor environmental condition, these changes will only become visible after the environmental conditions improve and this could be already too late. This paper is based on a different method of analysis: Fuzzy logic. We present an experiment that is designed to reproduce foggy conditions and discuss possible approaches for the analysis of observations. Peter Fisher states that fuzzy set theory is a method to deal with vague, poorly defined objects (Fisher, 1999). In the situation discussed here, the object itself is clearly defined but the environmental conditions, which affect the path that the light takes, is not. Thus we assume that fuzzy set theory is a promising approach to this kind of problem.

The remainder of the paper is organized as follows: In section 2 we give a brief introduction into fuzzy numbers and vectors and their analysis. In section 3 we present the experimental setup and show the resulting data. Section 4 contains some results and a discussion of two ideas for other approaches. Section 5 concludes the paper by providing directions for future work.

2. FUZZY NUMBERS AND FUZZY VECTORS

In statistics, an imprecise (fuzzy) number $x*$ is defined by its one dimensional characterizing function $\xi(\cdot)$, which is a real function of a real variable x with values between 0 and 1 (Menger, 1951, Zadeh, 1965 as quoted by Viertl, 2006a). The value represents the degree of membership. Fuzzy numbers are represented by δ-cuts for numerical treatment, where δ is a real number from the interval (0,1] (see Fig. 1). Each interval for a δ-cut contains any interval with for a larger value of δ.
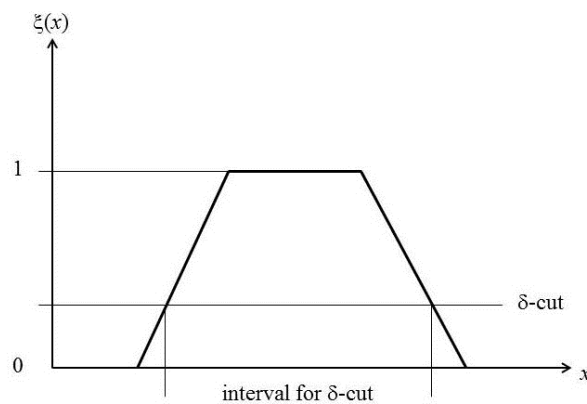

Figure 1. Fuzzy number and δ-cut

---

* Corresponding author

Data analysis then requires that the necessary functions are applied to the fuzzy numbers, i.e., $y^* = f(x_1^*, \ldots x_n^*)$. The result is again a fuzzy number and can be described by δ-cuts (Viertl, 2006b).

When moving from one dimensional fuzzy numbers to n-dimensional fuzzy vectors, different representations are possible (Liang et al., 2009). An n-dimensional fuzzy vector requires an n-dimensional characterizing function $\xi: R^n \rightarrow R$ in the interval (0,1]. A 2-dimensional fuzzy vector $\xi_2: R^2 \rightarrow R$ provides a membership value for each point in a plane. If the vector describes the horizontal extent of a mountain, then each point with $\xi_2=1$ is definitely part of the mountain and each point with $\xi_2=0$ is definitely not part of the mountain. In the area between these extremes, the point may be part of the mountain and the likelihood of the membership increases with the value of ξ_2. It is important to note, hat a 2-dimensional fuzzy vector is not a vector of two fuzzy numbers because this would only be true if the two parameters of the characterizing function can be separated to $\xi_2=\min(\xi(x),\xi(y))$.

Combination of fuzzy numbers or fuzzy vectors can then be done by using the δ-cuts. Fig. 2 shows an example of the combination of two fuzzy numbers with the minimum combination rule. The overlap of the δ-cuts is the result of the combination of the two fuzzy numbers. The same can be done with fuzzy vectors.
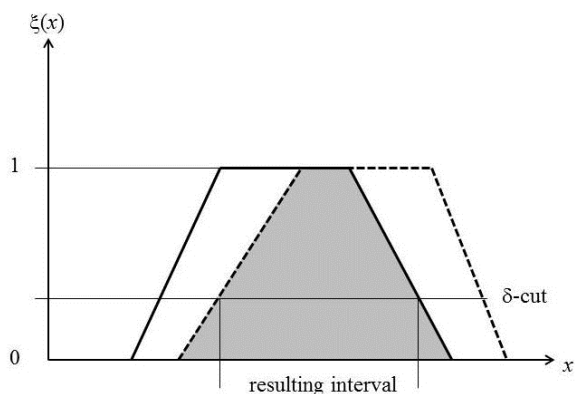

Figure 2. Combination of two fuzzy numbers

3. EXPERIMENTAL SETUP

The experiment was performed by Elsa Heer at TU München. The setup consists of a video theodolite, which was located at two different (known) positions (Fig. 3) and a light emitting target (Fig. 4). The results of a video theodolite observation are crisp and precise measurements and an image. Thus the system had to be disturbed for the experiment. This was done by setting focus and exposure of the video theodolite deliberately wrong. The result of the measurement does not have a clear maximum of intensity. A similar (and more realistic) effect could have been achieved by producing fog in the laboratory. However, this was not possible due to other experimental setups in the room that could have been affected. Fig. 5 shows one of the final images which is similar to an image is taken in a slightly foggy situation. Fig. 6 shows the image of a light source in foggy conditions. The foggy conditions show a similar pattern than the experiment but there are two differences: The area between light and darkness is bigger in Fig. 6 and the radiating pattern from Fig. 5 is missing. The advantage of the results

from the experiment is that the conditions can be easily reproduced.

The goal of this experiment is to compute the coordinates of the target using the directions determined by the theodolite. The theodolite was a Leica Geosystems TPS 1201with an attached CMOS camera.

Geometrically, the setup is a forward intersection. From two known points, the directions to a point with unknown coordinates are determined. The setup results in two lines that need to be intersected. Due to inevitable measurement deviations, the lines are skew lines and the intersection is typically computed by correcting the observations according to the least squares method. The size of the necessary corrections provides an estimate for the quality of the solution. In the setup the angle of intersection is 33 gon. This is not an optimal intersection angle but a realistic approach.

The basic setup thus consisted of four locations:
- 2 positions where the video theodolite was placed for the measurement (instrument positions),
- the position of the target, and
- a reference point to orient the video theodolite (this point is irrelevant for the further discussion because it is only used to align the measurements done with the video theoolite).

The reference point is necessary because theodolites measure horizontal directions with respect to an arbitrary starting direction. In order to use the observations for the determination of coordinates in a given reference frame, the observations need to be oriented, i.e., they need to be rotated such that the reference direction corresponds with the x-axis of the reference frame. This is done by observing the direction to a point where the coordinates are known (the reference point). In our setup the coordinates of all locations were also determined by classical measurements using a theodolite. Therefore, not only the coordinates of the instrument positions and the orientations are known, but also the coordinates of the target position. This guarantees that the quality of the final result can be determined.

Figure 3. Experimental setup: Instrument positions with the target in the background (pictures taken by Elsa Heer)

The intensity of the image can be used to determine the characterizing function. It is obvious, that the typical simplification for the shape of the characterizing function, which is shown in Figures 1 and 2, does not apply for this example. The real shape is more complex and less regular. The circular patterns may even lead to situations where the δ-cuts are not connected areas.

The result of the measurement is now modelled by a fuzzy vector in 3D space. In order to model it, two parameters are necessary to determine the direction of the vector. These parameters can either be modelled as vertical and horizontal

angles as observed from the projection centre of the theodolite or as horizontal and vertical displacement in a defined distance from the projection centre (the focal length of the optical system).
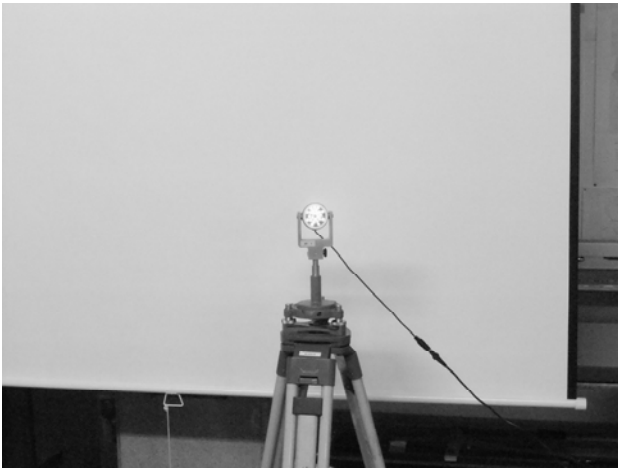

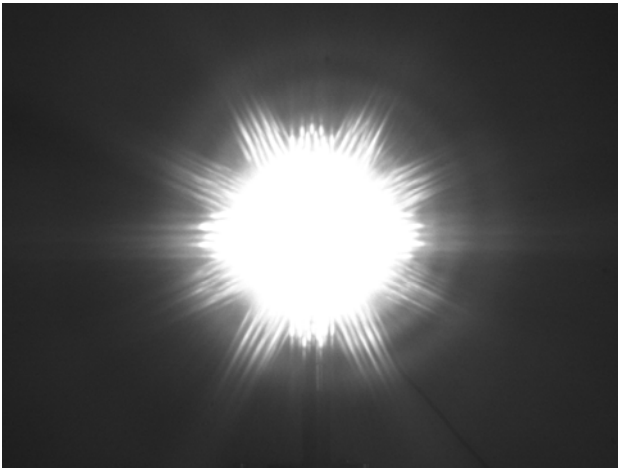Figure 4. Experimental setup: Target position (picture taken by Elsa Heer)

Figure 5. Image as taken by the video theodolite from point 2

Figure 6. Image of a light source in foggy conditions

4. PRELIMINARY RESULTS

In this section we discuss advantages and disadvantages of three methods that can be used to calculate the coordination of the light source.

4.1 Simple Method

In a first step, a simple method has been tested: In order to make decisions based on fuzzy information, the information is usually defuzzied (the area of the characterizing function is reduced to a point) and then further calculations are applied. The defuzzification can be based on the centroid of the characterizing function. Theoretically, this works if the uncertainty is purely random and thus the centroid is a reasonable approximation of the expectation of the distribution.

The δ-cuts of the characterizing function require a characterizing function in to range of [0,1]. Therefore the intensity of the images (e.g. Fig. 5) has to be normalized. One normalization result is included in Fig. 7. This characterizing function can then be used to compute the δ-cuts.
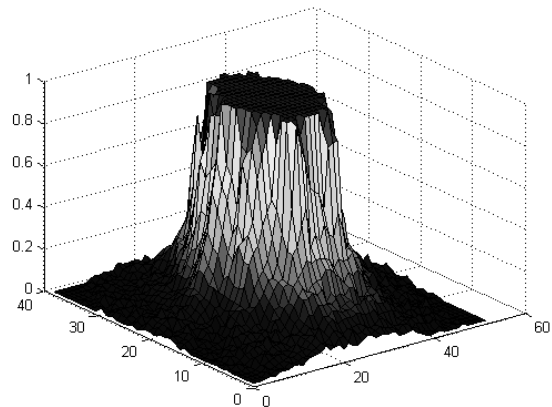

Figure 7. Characterizing function for the direction from point 2 (Heer, 2014)

An open question is, which δ-cuts should be used. In the following, the values 0.25, 0.5, 0.75, and 1 are used for δ. Each of these intersections results in a different area. The centroid of the area is then used to determine the line from the instrument positions to the target position. Once this is done for both instrument positions, the most likely intersections can be computed.

The intersections for the selected δ-cuts are summarized in Table 1. The x- and y-coordinates differ significantly from the results computed with the undisturbed observations (using the full precision of the theodolite) since the distance is short, the atmospheric conditions are controlled, and the theodolite is of high quality, any deviation of more than a few centimetres is a failure. The results made it obvious that this approach is not suitable for the analysis of fuzzy observations, because of differences of max. 1 meter in x direction and of max. 7 meters in y direction.

	x [m]	y [m]	z [m]
$\delta=0.25$	101.286	117.040	49.749
$\delta=0.5$	99.113	110.455	49.869
$\delta=0.75$	99.428	113.649	49.821
$\delta=1$	100.047	123.086	49.668
undisturbed result	101.303	117.754	49.643

Table 1. Estimated point coordinates for different δ-cuts and the accurate result from the undisturbed measurement

4.2 Cone Boundary Intersection Method

Another approach would be intersecting the cones defined by the δ-cuts. The centre of the optical system of the theodolite is known (typically by the coordinates of the point and the vertical offset of the theodolite). The angular readings of the theodolite determine the direction in which the telescope is pointing. Since the video camera is attached to the telescope and the mutual orientation is known, the orientation and position of the image in relation to the centre point is known. A δ-cut consists of pixels with value 1 (the interior of the vector) and 0 (the exterior of the vector). The boundary of the cone is determined by the boundaries between pixels with different values. Each instrument position provides a cone and the intersection of these two cones is the volume of space that agrees with both observations.

A property of the cone intersection is that the direction of highest uncertainty is recognizable. Fig. 8 shows that the intersection area has a small lateral extent and a much larger longitudinal extent.

A problem of the cone intersection is the complexity of the δ-cuts. As shown in Fig. 5, the data are not clearly structured. The rays with high intensity and the concentric circles may lead to disconnected areas of δ-cuts and therefore to complex volumes. Methods to determine the boundaries need to be carefully programmed to avoid numerical problems.
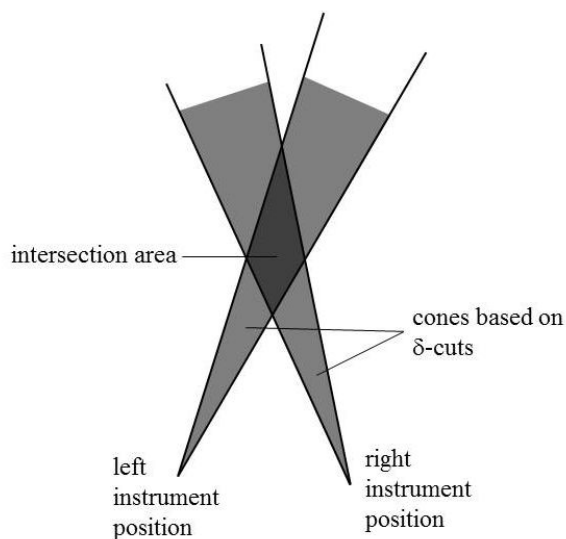
Another problem is the result itself. Surveyors typically work with random variables, which can be modelled as crisp values with an attached standard deviation. All computational algorithms have been optimized for this kind of input and it may require significant changes to include volumetric point information. Defuzzification of the resulting volume is not a good solution either. The volume is connected to a specific δ-cut. There is no information in the model providing probability of specific values. Typically, the volume will get smaller if the value for δ increases with the smallest volume for δ=1. However, only using δ=1 for the computation is dangerous because there may be situations where these cones do not overlap. Then lower numbers for δ have to be used in order to obtain a result. Therefore, in larger evaluations different δ-cuts need to be used for the whole project since it may be difficult to assess the optimal δ-cut at the beginning.

4.3 Grid Ray Intersection Method

Another method represents the cones as a bundle of rays. In order to calculate ray the characterizing function is stored as an image. Thus each pixel belonging to a specific δ-cut could be used to define a ray from the projection centre of the theodolite through the centre of the pixel. When using different δ-cuts, each ray belongs to a specific δ-cut. It may even be argued, that it belongs to any δ-cut with a δ lower than the value used to create the ray. If, for example, the pixel is part of the 0.5-cut but not part of the 0.75-cut, then it definitely is part of the 0.25-cut too, because the area of the 0.25-cut is at least as large of the area of the 0.5-cut and includes the 0.5-cut.

Fig. 9 illustrates the concept of the grid ray intersection method. Each cone is represented by a bundle of rays. These bundles are intersected and the resulting intersection points form the area of intersection, or in 3D the volume of intersection.
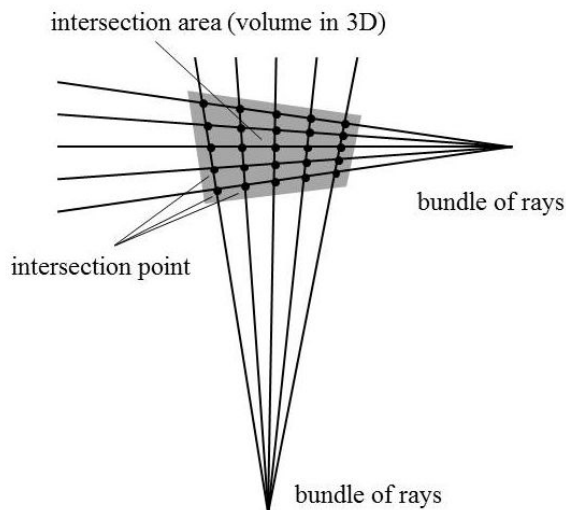
Figure 9. 2D-representation of the grid ray intersection method

In 3D, the rays will not intersect because they will be skew lines. This is true for any two lines from different bundles but some lines will pass each other at a close distance such that they can count as intersecting. It is necessary to identify these lines. It is possible to compute a midpoint between two arbitrary rays (the point in the middle of the closest connection between the

Figure 8. Concept of cone boundary intersection method; the position of the target is somewhere in the intersection area

two rays). The intersection volume can be generated by computing the midpoints for any two rays that 'almost intersect' (pass each other at a close distance). The criterion to determine that two rays are too far apart to be classified as 'almost intersecting' is critical for the evaluation. A too large space will be filled with points if the criterion is too weak and an insufficient number of points or no points at all will be found if the criterion is too strong. The criterion needs to be based on the approximate distance between projection centre and a reasonable starting value could be the lateral distance between two neighbouring rays from one bundle.

Since the basis for the bundle of rays is a regular raster, the following idea could be used: The lateral distance between two rays increases with the distance from the projection centre. The raster has an equally horizontal and vertical resolution and thus neighbouring points have a distance of d in the 4-neighbourhood and a maximum of $d \cdot sqrt(2)$. Half this maximum distance could be used to identify nearby rays. The identification of one pair of nearby rays (a,b) does not preclude the existence of another ray c that is also nearby the ray a. Thus all possible combinations have to be checked.

The result of the process is a set of points in space for each δ-cut. Since a δ-cut includes all rays for δ-cuts with a higher value for δ, the set of intersection points with shrink with an increasing value of δ. The space filled by the set of points is the δ-cut of the intersection.

The image size of the used CMOS camera is 1280x960 pixels. The δ-cut for δ=1contains approximately 160,000 pixel. Thus 25 billion combinations have to be checked for the smallest cone only. Assuming orthogonal rays and characterizing functions with a square extent of 400x400 pixels, then each ray from bundle 1 is nearby 400 pixels from the other ray. As a result, the point cloud consists of 64 million points. The limiting factor is not the large number of intersections or points, it is the fact that this is only a simple example and the numbers are already quite high.

5. CONCLUSIONS AND OPEN QUESTIONS

During the experiment and analysis it became obvious that the analysis of fuzzy observation data is not simple. The quality of the simple solution is unacceptable. Therefore, other strategies need to be developed and tested. Two strategies have been presented: The first approach uses the boundaries of the δ-cuts for the two observations. The intersection of these boundaries then defines the δ-cut of the final result. The second approach represents the δ-cuts by the contained points instead of their boundary and models the volume of the cones as a bundle of rays thorough these points.

Both solutions have advantages and disadvantages. The cone boundary intersection produces the boundary of the cone this provides a clear topological structure with an interior and an exterior for each δ-cut. However, computations may be demanding due to glancing intersections. The grid ray intersection, on the other hand, has a quite simple computational model. Apart from parallel rays there are not much computational pitfalls. The major problem in this approach is the huge number of computations and the size of the resulting data set. A possible solution could be the reduction of grid resolution. If each 10x10 square is represented by a ray then there are only 1,600 rays left in each bundle and only 2.56

million intersections need to be checked. However, applying this kind of strategy needs further investigation. Especially the dependency between reduction of the resolution and the deterioration of the quality needs to be analysed.

Another open question is the relation to the standard method to compute the result for the forward intersection. It is assumed that random error (noise) affects the observational data. The least squares approach provides best estimates for the noise and thus determines self-consistent measurement data, in this case two lines in 3D which intersect in a point. Is it possible to combine this approach with one of the strategies developed in this paper? Can we achieve better results if—in analogy to what happens to observation values in a least squares approach—we slightly move the cones?

The intersection of two cones has the largest volume if the axes of symmetry of the two cones meet. Thus larger overlapping areas do mean higher consensus between the symmetry axes but this may not necessarily be a sign of better quality in general.

ACKNOWLEDGMENT

REFERENCES

Fisher, P., 1999. Models of Uncertainty in Spatial Data. In: Geographical Information Systems: Principles, Techniques, Management and Applications, Longley, P, Goodchild, M., Maguire, D., and Rhind, D. (eds.), Vol. 1, John Wiley & Sons, Inc. Hoboken, New Jersey, pp. 191-205.

Ghilani, C.D., 2010. Adjustment Computations. 5th ed., John Wiley & Sons, Inc. Hoboken, New Jersey, 647 p.

Heer, E., 2014. Positionsbestimmung mittels unscharfer Information. Bachelor Thesis at TU Vienna, Department for Geodesy and Geoinformation, Vienna, Austria.

Liang, J., Navara, M. & Vetterlein, T., 2006. Different Representations of Fuzzy Vectors. In: *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Springer, Berlin Heidelberg, LNCS Vol. 5590, pp. 700-711.

Menger, K., 1951. Ensembles flous et fonctions aleatoires. Comptes Rendus de l'Academie de Sciences Paris, France.

Viertl, R., 2006a. Univariate statistical analysis with fuzzy data. *Computational Statistics & Data Analysis*, 51(1), pp. 133-147.

Viertl, R., 2006b. Description of Data Uncertainty and its Consequences for Data Analysis. Report SM-2006-4, TU Vienna, Vienna, Austria.

Zadeh, L., 1965. Fuzzy Sets. Information and Control, **8**, pp. 338-353.