

CONTOUR APPROXIMATION OF DATA AND THE HARMONIC MEAN

MARINA ARAV

Abstract. A contour approximation of data is a function capturing the data points in its lower level-sets. Desirable properties of contour approximation are posited, and shown to be satisfied uniquely (up to a multiplicative constant) by the weighted harmonic mean of distances to the cluster centers. This harmonic mean is the joint distance function used in probabilistic clustering, expressing the uncertainty of classification.

Mathematics subject classification (2000): 62–07, 26E60, 91C20, 47A64.

Key words and phrases: Harmonic mean, quasi-linear mean, clustering, contour approximation of data, distance functions.

REFERENCES

- [1] J. ACZÉL, *Lectures on Functional Equations and their Applications*, Academic Press, 1966.
- [2] W. N. ANDERSON JR., *Shorted operators*, SIAM J. Appl. Math., 20, 520–525, 1971.
- [3] W. N. ANDERSON JR. AND R. J. DUFFIN, *Series and parallel of matrices*, J. Math. Anal. Appl., 26, 576–594, 1969.
- [4] W. N. ANDERSON JR., T. D. MORLEY AND G. E. TRAPP, *Ladder networks, fixpoints, and the geometric mean*, Circuits, Systems, and Signal Processing, 2, 259–268, 1983.
- [5] W. N. ANDERSON JR. AND G. E. TRAPP, *Shorted operators II*, SIAM J. Appl. Math., 28, 60–71, 1975.
- [6] A. BEN-ISRAEL AND C. IYIGUN, *Probabilistic distance clustering*, J. of Classification, to appear.
- [7] K. R. DIXON AND J. A. CHAPMAN, *Harmonic mean measure of animal activity areas*, Ecology, 61, 1040–1044, 1980.
- [8] R. J. DUFFIN, *Network Models*, Mathematical Aspects of Electrical Network Analysis, SIAM–AMS Proceedings, Vol. III, 65–91, American Mathematical Society, 1971.
- [9] C. IYIGUN AND A. BEN-ISRAEL, *Probabilistic distance clustering with cluster size*, Probability in Engrg. and Info. Sci., to appear.
- [10] F. KUBO AND T. ANDO, *Means of positive linear operators*, Mathematische Annalen, 246, 205–224, 1980.
- [11] T. D. MORLEY, *Parallel summation, Maxwell’s principle and the infimum of projections*, J. Math. Anal. Appl., 70, 33–41, 1979.
- [12] M. TEBoulLE, *A unified continuous optimization framework to center-based clustering methods*, Journal of Machine Learning, 8, 65–102, 2007.