Discussion of 'Maximum likelihood estimation of a multi-dimensional log-concave density' by M. Cule, R. Samworth and M. Stewart

Guenther Walther[*]

Department of Statistics, 390 Serra Mall

Stanford University, Stanford, CA 94305, USA

gwalther@stanford.edu

May 25, 2010

I started looking at log-concave distributions when I was searching for an appropriate model for subpopulations of multivariate flow cytometry data about ten years ago. The use of log-concave distributions is appealing for this purpose since their unimodal character is commonly associated with a single component population. In addition, log-concave distributions possess a certain non-parametric flexibility that is helpful in many problems, but they can still be estimated without having to deal with a tuning parameter. When I worked out how to compute the MLE in the univariate case, I realized that the multivariate case would be much more daunting, requiring a more involved optimization algorithm and considerable computational overhead for the construction of multivariate tessellations. I considered the task to be too challenging and decided not to pursue it further beyond the univariate work I had done at that time.

Cule, Samworth and Stewart have shown in their paper how to compute the multivariate MLE using Shor's r-algorithm, and they provide an accompanying software package that implements their algorithm. I congratulate them on this work and I believe their paper will inspire much further research into the multivariate case. In particular, they show how by modifying the objective function for the maximum likelihood estimator, the problem becomes amenable to known, albeit slow, convex optimation algorithms. It is desirable to improve upon the computation times given in Table 1, especially for the higher dimensional cases. I expect that the groundwork that the paper lays in terms of the optimization problem will inspire new research into faster algorithms. Another

intriguing result is the outstanding performance of the MLE vis-a-vis other nonparametric methods as reported in their simulation study. These results provide a strong motivation to establish theoretical results about the finite sample and asymptotic performance of the MLE.