ESTIMATION OF FINITE MIXTURES USING THE EMPIRICAL CHARACTERISTIC FUNCTION

N86-14079

By

Charles Anderson, Associate Professor
Thomas Boullion, Professor
Department of Mathematics and Statistics
University of Southwestern Louisiana
Lafayette, Louisiana 70504

ABSTRACT

A problem which occurs in analyzing Landsat scenes is the problem of separating the components of a finite mixture of several distinct probability distributions. A review of the literature indicates this is a problem which occurs in many disciplines, such as engineering, biology, physiology and economics. Many approaches to this problem have appeared in the literature; however, most are very restrictive in their assumptions or have met with only a limited degree of success when applied to realistic situations.

We have been investigating a procedure which combines the "k-L procedure" of [Feurverger and McDunnough, 1981] with the "MAICE" procedure of [Akaike, 1974]. The feasibility of this approach is being investigated numerically via the development of a computer software package enabling a simulation study and comparison with other procedures.

Center Research Advisor: M. C. Trichel

INTRODUCTION

A Problem which occurs in many disciplines is that of separating the components of a probability distribution which is a finite mixture of several distinct distributions. See, for instance, [Yakowitz, 1970], [Bhattacharya, 1976], and [Day, 1969]. This problem is encountered in the Remote Sensing Research Branch of NASA Johnson Space Center in analyzing Landsat data.

A number of different approaches have been taken to resolve this problem, each enjoying a rather limited degree of success or being too restrictive to be widely applicable. Since the likelihood function corresponding to finite mixtures of normal distributions is unbounded, maximum likelihood estimation frequently breaks down in practice. The estimator which minimizes the sum of squares of differences between the theoretical and sample moment generating functions, given by [Quandt and Ramsey, 1978], seems to suffer from inefficiency and some arbitrariness in the choice of weights given to the moments. Estimating the mixing proportions of a mixture of known distributions [Bryant and Paulson, 1983], using the distance between characteristic functions is too restrictive, since it assumes the parameters in the component distributions are completely known.

A recent approach by [Heydorn and Basu, 1983] makes use of a constructive proof of a theorem of Caratheodory on a trigonometric moment problem, as discussed in [Grenander and Szego, 1958], to determine identifiable mixtures for certain special cases of families of distributions. When only sample data is available, this approach does not seem to be immediately applicable.

PROCEDURE AND JUSTIFICATION

The "k-L procedure" introduced by [Feurverger and McDunnough, 1981], refers generally to approximate maximum likelihood estimation based on the asymptotic distribution at k points of the empirical characteristic function (e.c.f.). Since the e.c.f. contains all the information in the sample, and for other reasons given later, it seems to be a promising technique. See Figure 1.

Let ξ be a column vector composed of the real and imaginary parts of the e.c.f. at points d, 2d,..., kd. The probability distribution of ξ is approximately multivariate normal, even for fairly small sample sizes, because of the smoothness and boundedness of the trigonometric functions. The covariance matrix $\Omega = E(\xi - E(\xi))\,(\xi - E(\xi))^T$ is determined by the values of the true characteristic function $\emptyset(t)$ at t=d, 2d,..., 2kd and can be estimated from the values of the e.c.f. at these points. Since this estimate $\hat{\Omega}$ is consistent, the following estimation criteria are asymptotically equivalent: (1) maximize the likelihood given ξ, (2) minimize $(\xi - E(\xi))^T \Omega^{-1}(\xi - E(\xi))$, (3) minimize $(\xi - E(\xi))^T \hat{\Omega}^{-1}(\xi - E(\xi))$.

The hypothesis H_O: $\emptyset(t) = \sum_{1}^{M} P_i\, \emptyset_i(t)\,|\Theta_{i_1}, \ldots)$ with certain of the parameters P_i, Θ_{i_j} specified, can be tested against an alternative hypothesis specifying the same form of the model but with parameters not all as specified by H_O, using the approximate chi-square distribution of $L = n(\xi - E(\xi))^T \Omega^{-1}(\xi - E(\xi))$. The hypothesis H_O is rejected if L is greater than the $(1-\alpha)$ point of its null distribution, which is approximately the chi-square with degrees of freedom equal to 2k minus the number of functionally independent unspecified parameters under H_O.

When there are several competing models, the MAICE procedure, introduced by [Akaike, 1974], selects the model which gives the minimum value of

AIC = (-2) log (maximum likelihood) +
2(number of independently adjusted
parameters within the model).

This procedure has been investigated by [Redner, Kitagawa, and Coberly, 1981], working directly with the mixed distributions. Our procedure applies the MAICE method to data reduced to a few carefully selected points of the e.c.f. Also, maximum likelihood is computed approximately using the approximate normality of the e.c.f.

FURTHER INVESTIGATIONS

Although the above procedure is based on sound theoretical arguments, no numerical results have appeared indicating its efficiency of implementation on a computer or its accuracy in determining the best model and estimates of the parameters. Thus, a software package is being developed which will enable us to implement and test the procedure. This work has led to numerical and theoretical investigations on optimally selecting the points t_i where the e.c.f. is evaluated. Other numerical problems are being investigated to determine efficient computational procedures and to increase the acccuracy of the computed values. The package is written in FORTRAN 77 and uses IMSL subroutines whenever possible. Basic components are (1) a very flexible subroutine (MIXSIM) to simulate data from any specified mixture of standard distributions, (2) an equally flexible subroutine (THEOCF) which computes the theoretical characteristic function for any specified mixture distribution, and (3) a subroutine (FITCF) which seeks parameter estimates, given a specified mixture model, to minimize the chi-square criterion L given above.

REFERENCES

1. Feurverger, A., and McDunnough, P., "On the Efficiency of Empirical Characteristic Function Procedures," J. R. Statist. Soc. B., 43:20-27, 1981.

2. Koutrouvelis, I. A., and Kellermeier, J., "A Goodness-of-Fit Test Based on the Empirical Characteristic Function when Parameters must be Estimated," J. R. Statist. Soc. B., 43:173-176, 1981.

3. Akaike, H., "A New Look at the Statistical Model Identification," IEEE Trans. Auto. Control, AC-19, No. 6, pp. 716-723, 1974.

4. Bhattacharya, C. G., "A Simple Method of Resolution of a Distribution into Gaussian Components," Biometrics, 23:115-135, 1967.

5. Day, N. E., "Estimating the Components of a Mixture of Normal Distributions," Biometrics, 56:463-474, 1969.

6. Yakowitz, S. J., "Unsupervised Learning and the Identification of Finite Mixtures," IEEE Transaction on Info. Theory, IT-16:330-338, 1970.

7. Quandt, R. E., and Ramsey, J. B., "Estimating Mixtures of Normal Distributions and Switching Regressions," JASA, 73# 364, 730-738. Comments 738-752, 1978.

8. Bryant, J. L., and Paulson, A. S., "Estimation of Mixing Proportions via Distance Between Characteristic Functions," Commun. Statist.-Theor. Meth., 12(9), pp. 1009-1029, 1983.

9. Feurverger, A., and Mureika, R. A., "The Empirical Characteristic Function and Its Applications," Ann. Statist., 5:86-97, 1977.

10. Redner, R. A., Kitagawa, G., and Coberly, W. A., "The Akaike Information Criterion and Its Application to Mixture Proportion Estimation," NASA-JSC Report SR-TI-04207, prepared for EOD under Contract NAS9-14689-95, 1981.

11. Heydorn, R. P., and Basu, R., "Estimating Location Parameters in a Mixture Model," NASA-JSC Internal Note, 1983.

12. Genander, U., and Szego, G., Toeplitz Forms and Their Applications, (Berkeley, University of California Press, 1958).

CHARACTERISTIC FUNCTIONS

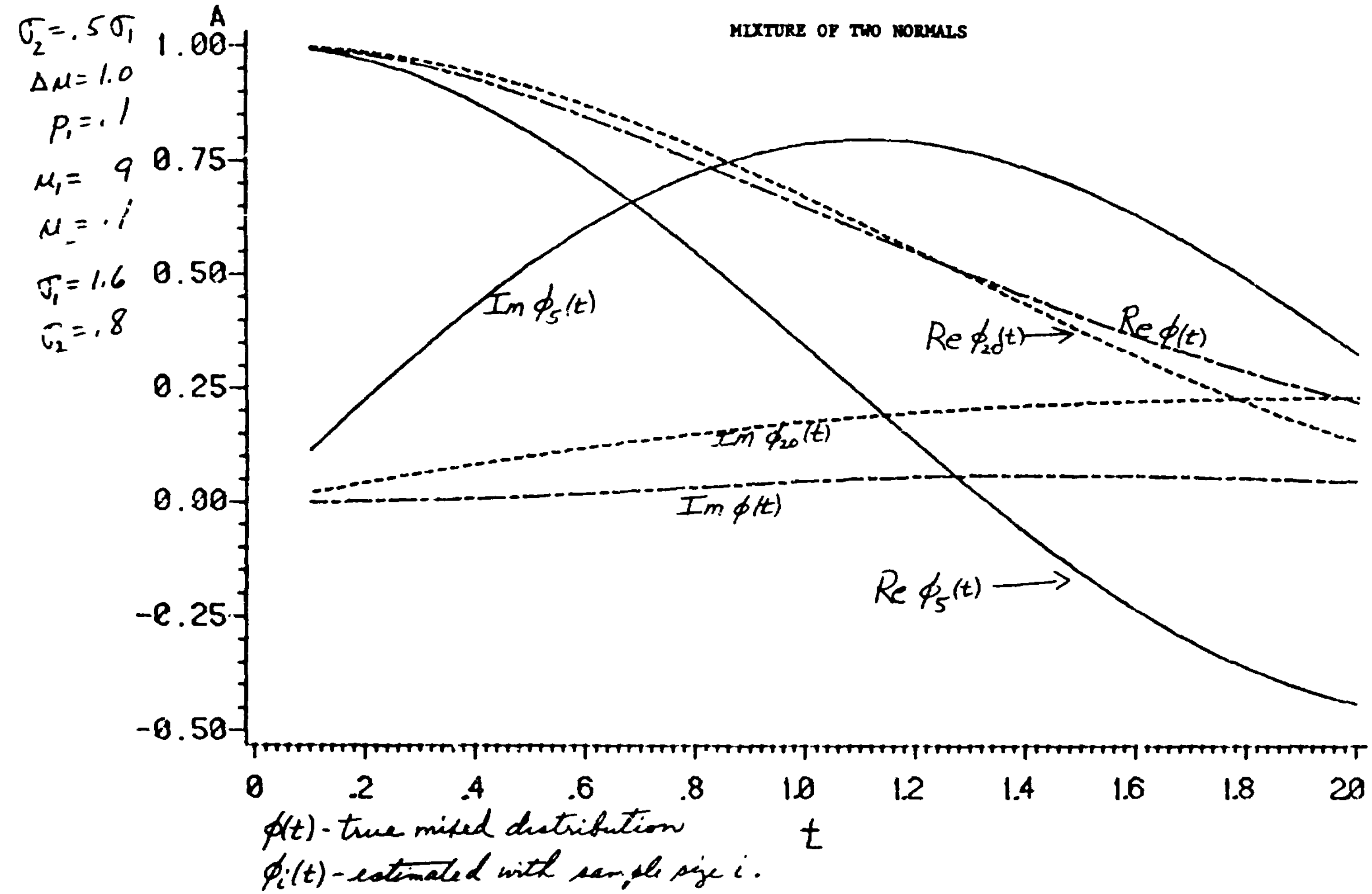

Figure 1