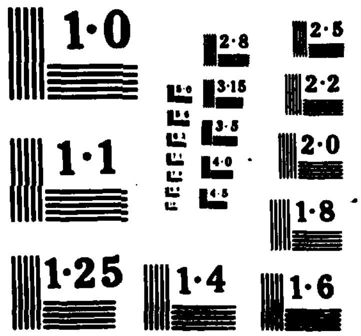
AD-A170 809 PATTERN RECOGNITION BASED ON SCALE INVARIANT DISCRIMINANT FUNCTIONS..(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS T M PUKKILA ET AL. APR 86 TR-86-09 AFOSR-TR-86-0575 F49620-85-C-0008 F/G 12/1

1/1

UNCLASSIFIED

NL


1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5

AFOSR-TR- 86-0575

(2)

AD-A170 809

ila and C. Radhakrishna Rao | F49620-85-(

ANIZATION NAME AND ADDRESS ltivariate Analysis Hall Pittsburgh, Pittsburgh, PA 15260	10. PROGRAM E AREA & WOF 6.1102F-
FICE NAME AND ADDRESS ice of Scientific Research the Air Force orce Base, DC 20332 NM	12. REPORT DA April 198 13. NUMBER OF 15
NCY NAME & ADDRESS(if different from Controlling Office) ame as NM	15. SECURITY C Unclassif

Approved for public release,
distribution unlimited

Center for Multivariate Analysis

University of Pittsburgh

DTIC FILE COPY

DTIC
ELECTE
AUG 1 2 1986
S D
B

86 8 164

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) [redacted]

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 86-0575	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Pattern Recognition Based on Scale Invariant Discriminant Functions		5. TYPE OF REPORT & PERIOD COVERED Technical Report - April1986
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Tarmo M. Pukkila and C. Radhakrishna Rao		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis, 515 Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 6.1102F-2304-A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research, Department of the Air Force, Bolling Air Force Base, DC 20332 NM		12. REPORT DATE April 1986
		13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) same as NM		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Angular Gaussian Distribution, Compositional Gaussian Distribution, Discriminant function, Scale free methods.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Some probability models for classifying individuals as belonging to one of two or more populations using scale invariant discriminant functions are considered. The investigation is motivated by practical situations where the observed data on an individual are in the form of ratios of some basic measurements or measurements scaled by an unknown non-negative number. The probability models are obtained by considering a p-vector random variable X with a known distribution and deriving the distribution of the random vector $Y = [G(X)]^{-1}X$, where G(X) is a non-negative measure of size such that $G(\lambda X) = \lambda G(X)$

DD FORM 1473, 1 JAN 73

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) for $\lambda>0$. Explicit expressions are obtained for the densities of what are called Angular Gaussian, Compositional Gaussian, Type 1 and Compositional Gaussian, Type 2 distributions.

PATTERN RECOGNITION BASED ON SCALE INVARIANT DISCRIMINANT FUNCTIONS*

by
Tarmo M. Pukkila
Department of Mathematical Sciences
University of Tampere
P.O. Box 607
SF-33101 Tampere, Finland
and
C. Radhakrishna Rao
Center for Multivariate Analysis
University of Pittsburgh
Pittsburgh, PA 15260

April 1986

Technical Report No. 86-09

Center for Multivariate Analysis
515 Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

*Research sponsored by the Air Force Office of Scientific Research (AFSC) under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

ABSTRACT

Some probability models for classifying individuals as belonging to one of two or more populations using scale invariant discriminant functions are considered. The investigation is motivated by practical situations where the observed data on an individual are in the form of ratios of some basic measurements or measurements scaled by an unknown non-negative number. The probability models are obtained by considering a p-vector random variable X with a known distribution and deriving the distribution of the random vector $Y = [G(X)]^{-1}X$, where G(X) is a non-negative measure of size such that $G(\lambda X) = \lambda G(X)$ for $\lambda>0$. Explicit expressions are obtained for the densities of what are called Angular Gaussian, Compositional Gaussian, Type 1 and Compositional Gaussian, Type 2 distributions.

Key words: Angular Gaussian Distribution, Compositional Gaussian Distribution, Discriminant function, Scale free methods.

1. INTRODUCTION

We consider the problem of classifying an individual as belonging to one of two or more populations using scale invariant discriminant functions. The investigation is motivated by practical situations where the observed data on an individual are in the form of ratios of some basic measurements or measurements scaled by an unknown non-negative number. In this paper we derive some probability models for applications to such data.

If $X' = (x_1,\cdots,x_p)$ is a vector of p basic measurements which may be known apart from a positive scaling factor, then we may consider transformed measurements

$$(y_1,\cdots,y_p)' = Y = [G(X)]^{-1}X \tag{1.1}$$

which are scale free if G is some non-negative measure of size such that $G(\lambda X) = \lambda G(X)$ for $\lambda > 0$. Some typical examples of G(X) are

$$G(X) = \|X\| = (\textstyle\sum x_i^2)^{\frac{1}{2}}, \tag{1.2}$$

$$= |\textstyle\sum x_i|. \tag{1.3}$$

We call the corresponding transformed variables $Y = X/\|X\|$, $X/|\sum x_i|$ as directional, and compositional data respectively. We note that the term compositional data is usually applied to a set of non-negative proportions (see Aitchison (1985)), but our definition is more flexible. However, we refer to $Y = X/|\sum x_i|$ as compositional data of type 1 and $Y = X/\sum x_i$ as of type 2, even when x_i are not non-negative.

It is also interesting to note that when we have compositional data with non-negative proportions, $(y_1,\cdots,y_p)$ such that $\sum y_i = 1$, then we may transform them into directional data by considering $(\sqrt{y_1},\cdots,\sqrt{y_p})$ and use appropriate probability models for directional data (with non-negative components) for statistical analysis as suggested by Stephens (1982).

One way of generating probability models for directional and compositional data is to consider a probability distribution for the basic measurements X and then derive the induced distribution for $Y = [G(X)]^{-1}X$. In this paper we assume that $X \sim N_p(\mu,\Sigma)$, i.e., as p-variate normal with mean vector μ and variance-covariance matrix Σ, and derive the distribution of Y for different size functions G.

Once an appropriate probability model is chosen, the problem of discrimination can be handled in the usual way.

We also comment on non-parametric methods for estimation of density for directional and compositional data.

2. CLASSES OF DISTRIBUTIONS FOR DIRECTIONAL DATA

2.1 Angular Gaussian Distribution (AGD)

Let $X \sim N_p(\mu, \Sigma)$ and define $Y = R^{-1}X$ where $R = ||X|| = (X'X)^{1/2}$, so that Y is the vector of direction cosines with the condition $Y'Y = 1$. The marginal distribution of Y on the p-dimensional unit sphere Ω_p is called the AGD (Angular Gaussian Distribution). For the special case when $\Sigma = \sigma^2 I$, Bingham obtained the distribution of Y in the form of an infinite series (see Watson (1983), p. 226). In this section, we obtain the distribution of Y in the general case in a closed form involving a finite number of terms.

Consider a polar transformation from X to (R, θ), where θ is a vector of p-1 angles in which case the Jacobian is of the form

$$\frac{D(X)}{D(R,\theta)} = R^{p-1} f(\theta). \tag{2.1}$$

The transformed p.d.f. (probability density function) is

$$|2\pi\Sigma|^{-\frac{1}{2}} R^{p-1} f(\theta) \exp[-2^{-1}(R^2 Q_3 - 2RQ_2 + Q_1)] \tag{2.2}$$

where

$$Q_3 = Y'\Sigma^{-1}Y,\ Q_2 = \mu'\Sigma^{-1}Y,\ Q_1 = \mu'\Sigma^{-1}\mu$$

and Y is a function of θ only. Changing from R to $r = R\sqrt{Q_3}$, the p.d.f. (2.2) transforms to

$$|2\pi\Sigma|^{-1/2}Q_3^{-p/2}f(\theta)\exp[-2^{-1}(Q_1-Q_2^2Q_3^{-1})]$$
$$\times\ r^{p-1}\exp[-2^{-1}(r-Q_2Q_3^{-1/2})^2]. \tag{2.3}$$

Integrating out for r from 0 to ∞, the p.d.f. at Y w.r.t. to the surface element $d\omega_p$ on Ω_p is

$$p(Y|\mu,\Sigma) = |2\pi\Sigma|^{-\frac{1}{2}}Q_3^{-p/2}I_p(Q_2Q_3^{-\frac{1}{2}})\ \exp[-2^{-1}(Q_1-Q_2^2Q_3^{-1})] \tag{2.4}$$

where

$$I_p(\alpha) = \int_0^\infty r^{p-1}\exp[-2^{-1}(r-\alpha)^2]dr. \tag{2.5}$$

The function I_p satisfies the recurrence relation

$$I_p(\alpha) = (p-2)\ I_{p-2}(\alpha) + \alpha\ I_{p-1}(\alpha)\ \text{for } p>2 \tag{2.6}$$

with the initial values

$$I_2(\alpha) = e^{-\alpha^2/2} + \alpha I_1(\alpha),\quad I_1(\alpha) = \sqrt{2\pi}\ \Phi(\alpha)$$

where $\Phi(\alpha)$ is the distribution function of $N_1(0,1)$.

It may be noted that the p.d.f. (2.4) remains unchanged if μ and Σ are replaced by $a\mu$ and $a^2\Sigma$ for any $a>0$. However, we can make the parameters unique by imposing the condition $||\mu|| = 1$.

2.2 Longevin Distribution

We can generate other distributions for Y from (2.2) and (2.3) as suggested by Fisher (1953) by considering conditonal instead of marginal distributions. Thus, from the expression (2.3), the conditional p.d.f. of Y on Ω_p given r = 1 is

$$\begin{aligned} &\text{const. } Q_3^{-p/2}\exp(Q_2 Q_3^{-1/2}) \\ &= \text{const. } (Y'\Sigma^{-1}Y)^{-p/2}\exp(\mu'\Sigma^{-1}Y/\sqrt{Y'\Sigma^{-1}Y}) \end{aligned} \tag{2.7}$$

where we may impose the restriction $||\mu|| = 1$. When $\Sigma = \sigma^2 I$ we have the Longevin (1905) – von Mises (1918) – Fisher (1953) distribution

$$\text{const. } \exp \kappa\mu'Y, \tag{2.8}$$

on the surface of a p-dimensional sphere.

From the expression (2.2), we find that the conditional p.d.f. of Y on Ω_p given R = 1 with respect to $d\omega_p$ is

$$\text{const. } \exp[-2^{-1}(Y-\mu)'\Sigma^{-1}(Y-\mu)] \tag{2.9}$$

where $||\mu|| = 1$.

We add two other classes of distributions found to be useful in practical applications as possible models for directional data:

<u>Scheideggar-Watson p.d.f.</u> $[b_p(\kappa)]^{-1}\exp(\kappa'Y)^2$. (2.10)

<u>Bingham p.d.f.</u> $[b(k)]^{-1}\exp Y'KY$, where K is p×p symmetric matrix. (2.11)

2.3 Estimation of Parameters

The model (2.4), which is the angular normal distribution, can be used to construct scale invariant discriminant functions provided the parameters μ and Σ are known. If they are unknown we may have to estimate them from past observations $Y_1,\cdots,Y_n$ on Y. Using the density function (2.4), the likelihood based on past data is

$$\prod_{i=1}^{n} p(Y_i|\mu,\Sigma)$$

with the restriction $\|\mu\| = 1$. The method of maximum likelihood for estimation of parameters can be implemented without much difficulty since the derivatives of all the expressions involved in (2.4) with respect to μ and Σ can be easily evaluated. However, there are too many parameters to be estimated and a very large sample may be necessary to obtain reasonably good estimators.

We may consider an alternative method by considering the marginal bivariate distributions of Y, where $Y = (x_1/\|x\|,\cdots,x_n/\|x\|)$ and $X \approx N_p(\mu,\Sigma)$. If y_1 and y_2 are the first two components of Y, then it is easily seen that

$$p_\alpha = P(y_1 \leq \alpha y_2) = P(x_1-\alpha x_2 \leq 0)$$

$$= \Phi[(\alpha\mu_2-\mu_1)/(\alpha^2\sigma_{22}-2\alpha\sigma_{12}+\sigma_{11})^{\frac{1}{2}}] \quad (2.13)$$

where Φ is the distribution function of $N_1(0,1)$. If we have a sample of size n on Y with the first two components (y_{1i},y_{2i}), $i = 1,\cdots,n$, we can estimate p_α for any given α by

$$\hat{p}_\alpha = \text{proportion of } i\text{'s such that } y_{1i} \leq \alpha y_{2i} \tag{2.14}$$

Then we have the observational equations

$$\Phi[(\alpha\mu_2-\mu_1)/(\alpha^2\sigma_{22}-2\alpha\sigma_{12}+\sigma_{11})^{1/2}] = \hat{p}_\alpha \ \forall\ \alpha. \tag{2.15}$$

or

$$(\alpha\mu_2-\mu_1)^2 = [\Phi^{-1}(\hat{p}_\alpha)]^2(\alpha^2\sigma_{22}-2\alpha\sigma_{12}+\sigma_{11}) \ \forall\ \alpha. \tag{2.16}$$

$p(p-1)/2$ families of equations of the kind (2.15) or (2.16) are available involving all the elements of μ and Σ by considering every pair of components in Y. From the equations (2.15) or (2.16) it is clear that only ratios of the parameters can be estimated. They can be made unique by using a restriction like $\|\mu\| = 1$. An appropriate method may be used to combine the equations (2.15) or (2.16) to produce the requisite number of consistent equations to estimate the parameters.

We describe one of the methods. First, we note that by smoothing $\hat{p}_\alpha$ in (2.15), we can estimate α_{-1}, α_0, α_1 such that $p_{\alpha_1} = 1/4$, $p_{\alpha_0} = 1/2$ and $p_{\alpha_1} = 3/4$. Writing

$$\Phi^{-1}(1/4) = q_{-1}\,,\quad \Phi^{-1}(1/2) = 0\,,\quad \Phi^{-1}(3/4) = q_1 \tag{2.17}$$

the equations (2.16) for α_{-1}, α_0, α_1 can be written as

$$(\alpha_0\hat{\mu}_2-\hat{\mu}_1) = 0 \tag{2.18}$$

$$(\alpha_s\hat{\mu}_2-\hat{\mu}_1)^2 = q_s^2\ (\alpha_s^2\sigma_{22}-2\alpha_s\sigma_{12}+\sigma_{11}),\quad s = -1, 1. \tag{2.19}$$

There are p equations of the kind (2.18) obtained by considering all pairs of the components of Y. They yield estimates of the ratios of $\mu_1,\dots,\mu_p$, which can be standardised to satisfy the restriction $||\mu|| = 1$. Then we have p(p−1) equations of the type (2.19) involving the p(p+1)/2 parameters in Σ. Observing that the equations are linear in σ_{ij}, we may combine them by least squares method to produce p(p+1)/2 equations by solving which we obtain the estimates of σ_{ij}.

The estimates obtained by the above method may still require large samples. Other methods of combining the equations (2.15) or (2.16) have to be explored.

3. CLASSES OF DISTRIBUTIONS FOR COMPOSITONAL DATA

3.1 Compositional Gaussian Distribution, Type 1, [CGD(1)]

Let $X \sim N_p(\mu, \Sigma)$ and define $Y = |\Sigma x_i|^{-1}X$ so that $|\Sigma Y_i| = 1$. We call the distribution of Y on the set

$$S = \{Y : |\Sigma y_i| = 1\} \tag{3.1}$$

the Compositional Gaussian Distribution, Type 1. We distinguish two sets

$$S_+ = \{Y : \Sigma y_i = 1\},\ S_- = \{Y : \Sigma y_i = -1\} \tag{3.2}$$

and note that $S_+ \cup S_- = S$ and

$$P(Y \in S_+) = P(\Sigma x_i > 0) \text{ and } P(Y \in S_-) = P(\Sigma x_i < 0). \tag{3.3}$$

In S_+ we consider the transformation

$$x_i = Ry_i,\ i = 1,\dots,p-1$$

$$x_p = R(1-y_1-\dots-y_{p-1}) = Ry_p \tag{3.4}$$

and in S_-,

$$x_i = -Ry_i, \; i = 1,\cdots,p-1$$

$$x_p = R(1+y_1+\cdots+y_{p-1}) = -Ry_p. \qquad (3.5)$$

The Jacobian of the transformation in either case is

$$\frac{D(x_1,\cdots,x_p)}{D(R,y_1,\cdots,y_{p-1})} = R^{p-1} \qquad (3.6)$$

The p.d.f. of X in S_+ transforms to

$$(2\pi)^{-p/2}|\Sigma|^{-\frac{1}{2}}R^{p-1}\exp[-2^{-1}(Q_3R^2-2Q_2R+Q_1)] \qquad (3.7)$$

where

$$Q_3 = Y'\Sigma^{-1}Y, \; Q_2 = \mu'\Sigma^{-1}Y, \; Q_1 = \mu'\Sigma^{-1}\mu.$$

Making the transformation $r = RQ_3^{1/2}$, the expression (3.6) changes to

$$(2\pi)^{-p/2}|\Sigma|^{-1/2}Q_3^{-p/2}\exp[-2^{-1}(Q_1-Q_2^2Q_3^{-1})]$$

$$\times\, r^{p-1}\exp[-2^{-1}(r-Q_2Q_3^{-1/2})^2]. \qquad (3.8)$$

Integrating out with respect to r from 0 to ∞, the p.d.f. of Y in S_+ with respect to the volume element $dy_1 \ldots dy_{p-1}$ is

$$(2\pi)^{-p/2}|\Sigma|^{-\frac{1}{2}}Q_3^{-p/2}I_p(Q_2Q_3^{-\frac{1}{2}})\exp[-2^{-1}(Q_1-Q_2^2Q_3^{-1})] \qquad (3.9)$$

where $I_p(\alpha)$ is as defined in (2.5).

In S_{-1}, under the transformation (3.4), the expression corresponding to (3.8) is

$$(2\pi)^{-p/2}|\Sigma|^{-1/2}Q_3^{-p/2}\exp[-2^{-1}(Q_1-Q_2^2Q_3^{-1})]$$

$$\times\ |r|^{p-1}\exp[-2^{-1}(r+Q_2Q_3^{-1/2})^2]. \qquad (3.10)$$

Integrating out with respect to r from $-\infty$ to 0, we obtain the same expression as in (3.9) for the pdf of Y in S_-.

It may be noted that in the expression (3.9), we can impose a suitable restriction on μ_i to make the parameters identifiable.

3.2 Compositional Gaussian Distribution, Type 2, [CGD(2)]

Let $X \sim N_p(\mu,\Sigma)$ and define $Y = (\Sigma x_i)^{-1}X$. For this we consider the transformation

$$x_i = Ry_i,\ i = 1,\cdots,p-1$$

$$x_p = R(1-y_1-\cdots-y_{p-1}) = Ry_p \qquad (3.11)$$

so that $\Sigma y_i = 1$. We define the marginal distribution of Y on the simplex

$$S = \{Y : \Sigma y_i = 1\}$$

as the Compositional Gaussian Distribution, Type 2. Making the transformation (3.11), proceeding as in Section 3.1 and integrating the expression corresponding to (3.8) with respect to r from $-\infty$ to ∞ we obtain the pdf of Y with respect to the volume element $dy_1 \ldots dy_{p-1}$ as

$$(2\pi)^{-\frac{1}{2}}|\Sigma|^{-\frac{1}{2}}Q_3^{-p/2}[I_p(Q_2Q_3^{-\frac{1}{2}}) + I_p(-Q_2Q_3^{-\frac{1}{2}})]\exp[-2^{-1}(Q_1-Q_2^2Q_3^{-1})] \quad (3.12)$$

where Q_1, Q_2 and Q_3 are as in (3.7) and I_p is as defined in (2.5).

As in the other cases, the pdf remains unchanged if μ and Σ are replaced by $a\,\mu$ and $a^2\Sigma$ respectively for any scalar $a > 0$.

As in (2.7), the conditional distribution of Y given $r = \kappa$ (a constant) is

$$\text{const } Q_3^{-p/2}\exp(\kappa Q_2 Q_3^{-1/2}) \quad (3.13)$$

which could be used as a probability model for compositional data.

If we define $Y = x_p^{-1}X$, then the distribution of $y_1,\cdots,y_{p-1}$ is the same as in (3.13). In the computation of Q_2 and Q_3, we substitute the value 1 for y_p. A natural way of normalizing the parameters μ and Σ is to consider

$$|\mu_p|^{-1}\mu \text{ and } \mu_p^{-2}\,\Sigma. \quad (3.14)$$

3.3 Logistic Gaussian and Related Distributions

Let Y be a vector of non-negative components $y_1,\cdots,y_p$ such that $\sum y_i = 1$. Then, one possible model which has been studied in detail is the logistic Gaussian distribution which assumes that

$$X' = (x_1,\cdots,x_{p-1}) = (\log y_1y_p^{-1},\cdots,\log y_{p-1}y_p^{-1}) \quad (3.15)$$

has a (p-1)-variate Gaussian distribution (see Aitchison and Shen (1980), Aitchison (1982)). In such a case the pdf of Y can be written in the form

$$(2\pi|\Sigma|)^{-\frac{1}{2}}(y_1\cdots y_p)^{-1}\exp[-2^{-1}(X-\mu)'\Sigma^{-1}(X-\mu)] \qquad (3.16)$$

where X can be expressed in terms of Y as in (3.15).

In building a model for Y we could have used other transformations from the basic (p−1) dimensional Gaussian variable X such as

$$x_i = \lambda^{-1}[(y_i/y_p)^{\lambda}-1],\ i = 1,\cdots,p-1 \qquad (3.17)$$

which is the Box and Cox (1964) transformation, or more generally any appropriate transformation

$$x_i = h_i(Y),\ i = 1,\cdots,p-1 \qquad (3.18)$$

suggested by data.

A well-known distribution for compositional data with non-negative proportions is the Dirichlet class $D_p(\beta)$ with the typical density function

$$[\Delta(\beta)]^{-1}y_1^{\beta_1-1}\cdots y_p^{\beta_p-1} \qquad (3.19)$$

where

$$\Delta(\beta) = \Gamma(\beta_1)\cdots\Gamma(\beta_p)/\Gamma(\beta_1+\cdots+\beta_p).$$

Aitchison (1985) considered a mixture of a Dirichlet and a logistic Gaussian distributions, but imposing some relationship between the parameters μ and β to reduce the number of free parameters in the model. He also provided a computational procedure for obtaining the maximum likelihood estimators of the parameters in such a mixture of distributions.

4. ESTIMATION OF THE DENSITY FUNCTION

Let $Y_1,\cdots,Y_n$ be independent observations on a random variable Y defined on Ω_p, the p-dimensional unit sphere. If a suitable model for the distribution of Y is not available, we may use non-parametric methods and estimate its p.d.f. based on $Y_1,\cdots,Y_n$. For this purpose we define a window function defined on Ω_p, which is indexed by two parameters x and θ, $x \in \Omega_p$ and $0 \le \theta \le \pi/2$,

$$\phi_{x,\theta}(Y) = \begin{cases} 1 \text{ if } x'Y \ge \cos\theta, \\ 0 \text{ otherwise.} \end{cases} \tag{4.1}$$

The set of points Y satisfying the first equation in (4.1) defines a cup on Ω_p with x as a central point, whose area is

$$a(\theta) = \frac{2\pi^{(p-1)/2}}{\Gamma(\frac{p-1}{2})} \int_0^\theta \sin\psi \, d\,\psi. \tag{4.2}$$

The number of points falling on this cup is

$$\sum_1^n \phi_{x,\theta}(Y_i). \tag{4.3}$$

By choosing a small value of $\theta = \theta_n$ an estimate of the p.d.f. of Y at x may be obtained as

$$p_n(x) = n^{-1} \sum_1^n [a(\theta_n)]^{-1} \phi_{x,\theta_n}(Y_i). \tag{4.4}$$

More generally we could use any suitable p.d.f. on Ω_p as a window function. In particular we suggest the use of the Longevin density (2.8)

$$c(\kappa_n)\exp(x'Y/\kappa_n) \tag{4.5}$$

and estimate the p.d.f. of Y as

$$p_n(x) = n^{-1}\sum_{1}^{n} c(\kappa_n) \exp(x'Y_i/\kappa_n). \tag{4.6}$$

We can choose κ_n by the method of Hebbema et al (1974) as the value κ at which the pseudo-likelihood

$$\prod_{i=1}^{n} (n-1)^{-1} \sum_{j\neq i} c(\kappa) \exp(Y_i'Y_j/\kappa) \tag{4.7}$$

is maximized. Further work on density estimation will be reported elsewhere.

5. REFERENCES

1. Aitchison, J. (1982). The statistical analysis of compositional data (with discussion). J.R. Statist. Soc. B, 44, 139-177.

2. Aitchison, J. (1985). A general class of distributions on the simplex. J.R. Statist. Soc. B, 47, 136-146.

3. Aitchison, J. and Shen, S.M. (1980). Logistic normal distributions: some properties and uses. Biometrika 67, 261-272.

4. Box, G.E.P. and Cox, D.R. (1964). The analysis of transformations (with discussion). J.R. Statist. Soc. B, 26, 211-252.

5. Fisher, R.A. (1953). Dispersion on a sphere. Proc. Roy. Soc. Lond. A217, 295-305.

6. Hebbema, J.D.F., Hermans, J. and van den Brock, K. (1974). A stepwise discriminant analysis program using density estimation. In Compstat 1974.

7. Langevin, P. (1905). Magnetisme et theorie des electrons. Ann. de Chim. et de Phys. 5, 70-127.

8. Mardia, K.V. (1972). Statistics of Directional Data. Academic Press, New York.

9. Rao, C. Radhakrishna (1973). Linear Statistical Inference and its Applications. Wiley, New York.

10. Stephens, M.A. (1980). Use of the von Mises distribution to analyse continuous proportions. Biometrika, 69, 197-203.

11. Von Mises, R. (1918). Uber die "Ganzzahligkeit" der Atomgewicht und Verwandte Fragen. Physikal. Z. 19, 490-500.

12. Watson, G.S. (1983) Statistics on Spheres. Wiley Interscience, New York.

END
DTIC
9-86