ON THE UNIMODALITY AND THE BIMODALITY OF THE MULTIVARIATE GAUSSIAN MIXTURES OF THE TWO COMPONENTS

N.N. Aprauseva[1], S.V. Sorokin[2]

Dorodnitsyn Computing centre центр of RAC, Russia, 119333, Moscow, Vavilov st., 40
[1] (495) 135.40.98 plat@ccas.ru
[2] (495) 135.14.98, www2008@ccas.ru

In the artical several sufficient conditions of the unimodality and necessary condition of the bimodality are formulated for the multivariate Gaussian mixtures of two components with the equal covariance matrixes Σ and with various vectors of expectation values μ_i, i = 1, 2. The results are received by the generalization of the theorems for an one-dimensional case.

Introduction

Wide use of Gaussian mixtures in various areas of a science and a practice [3, 6, 8] needs the definition of the number of its modes depending on the values of distribution parameters. In general the question on a number of the modes of the Gaussian mixture is not solved. In the article the solution of this problem is presented for the multivariate mixture of two components, with the equal covariance matrixes Σ and with the various vectors of the expectation values μ_i, i = 1, 2. Results are received by the generalizing the theorems for an one-dimensional case [4].

The Statement of the Problem and the Solution Methods

The probability density of such a mixture is expressed by the formula

$$f(X,\theta) = \left(|\Sigma|(2\pi)^p\right)^{-\frac{1}{2}} \cdot \sum_{i=1}^{2} \pi_i exp\left(-\frac{1}{2}(X-\mu_i)\Sigma^{-1}(X-\mu_i)\right) \qquad (1)$$

$$\theta = (\mu_1, \mu_2, \pi_1, \pi_2, \Sigma),$$

$X=(x_1, x_2, \ldots, x_p)$, $X \in R^p$, $p \geq 2$, Σ is a covariance matrix of i-th component $|\Sigma|>0$, μ_i is a vector of the expectation value i-th component, π_i is a priori probability, $\pi_i \in (0, 1)$, $\pi_1 + \pi_2 = 1$, X is a row-vector, X' is a column-vector.

Since function (1) and a Mahalanobis distance.

$$\rho^2 = (\mu_2-\mu_1)\,\Sigma^{-1}\,(\mu_2-\mu_1)'$$

are the invariants in relation to the linear orthogonal transformation, for simplification of proofs we shall go from the initial system of coordinates to the system of the principal components [1, 2]. Then

$$Y = X B, \quad v_i = \mu_i B, \quad i = 1, 2,$$
$$\mathrm{cov}(Y', Y) = \mathrm{cov}(B'X', XB) = B'\,\Sigma\,B = \Lambda,$$

where B is an orthogonal matrix of the eigenvectors of a matrix Σ, Λ — a diagonal matrix. In the new system of coordinates the function (1) becomes

$$f(Y, \theta^*) = \left(|\Lambda|(2\pi)^p\right)^{-\frac{1}{2}} \sum_{i=1}^{2} \pi_i exp\left(-\frac{1}{2}\sum_{s=1}^{p}\frac{(y_s - v_{is})^2}{\lambda_s}\right), \qquad (2)$$

$$\theta^* = (v_1, v_2, \pi_1, \pi_2, \Lambda),$$

The mode of the smooth function $f(Y, \theta^*)$ is its critical point (CP), i.e. its coordinates satisfy the system of the equations [7]

$$\frac{\partial}{\partial y_i} f(Y, \theta^*) = 0, \quad i = 1, 2, \ldots, p.$$

This system of the equations after differentiation of the function (2) with respect to y_i is reduced to the system

$$\pi_1 e^{K_1}(y_i - v_{1i}) + \pi_2 e^{K_2}(y_i - v_{2i}) = 0, \quad i = 1, 2, \dots, p, \qquad (3)$$
$$K_s = -\frac{1}{2}\sum_{i=1}^{p}\frac{(y_i - v_{si})^2}{\lambda_i}, \quad s = 1, 2.$$

We will present each equation of the system (3) in the form of

$$y_i - v_{1i} = -\frac{\pi_2}{\pi_1}(y_i - v_{2i})e^{K_2 - K_1}, \quad i = 1, 2, \dots, p, \qquad (4)$$

Since $v_1 \neq v_2$ then can put $\mu_{11} \neq \mu_{21}$. Then $y_1 = v_{11}$, $y_1 = v_{21}$ are not the roots of the first equation of the system (4), and each equation of this system at $i = 2, 3, \dots, p$ is possible to divide by the first equation. After such operations we have a system of the equations

$$\frac{y_i - v_{1i}}{y_1 - v_{11}} = \frac{y_i - v_{2i}}{y_1 - v_{21}}(y_i - v_{2i})e^{K_2 - K_1}, \quad i = 2, 3, \dots, p,$$

equivalent to the system

$$(v_{1i} - v_{2i})y_1 + (v_{2i} - v_{11})y_i + v_{11}v_{2i} - v_{1i}v_{21} = 0, \quad i = 2, 3, \dots, p, \qquad (5)$$

Each equation of the system (5) is an equation of the hyperplane which is collinear to the coordinate axis 0z and passing through one CP of the surfaces (2). From the equations (5) we have

$$y_i = \frac{v_{21}v_{1i} - v_{11}v_{2i} + (v_{2i}v_{1i})y_1}{v_{21} - v_{11}}, \quad i = 2, 3, \dots, p, \qquad (6)$$

Substituting the expressions (6) to the formula (2) we will receive

$$f(Y, \theta^*) = (|\Lambda|(2\pi)^p)^{-\frac{1}{2}} \cdot \left[\sum_{s=1}^{2}\pi_s e^{-\frac{(y_1 - v_{s1})^2\rho^2}{2(v_{21} - v_{11})^2}}\right], \qquad (7)$$

where ρ^2 is a Mahalanobis distance in the system of principal components

$$\rho^2 = \sum_{i=1}^{p}\frac{(v_{2i} - v_{1i})^2}{\lambda_i}.$$

The function (7) describes the plane curve which is passing through any CP of the surface (2) and received by the hyperplane set (5) section of this surface. Obviously, it is possible to receive as many parametrizations of this curve, as available inequalities $v_{1i} \neq v_{2i}$, $i \in \{1, 2, \dots, p\}$.

At p=1 the probability density of the investigated mixture is presented in the form of

$$f(y, \theta_1) = (\sigma^2 2\pi)^{-\frac{1}{2}} \cdot \left[\sum_{s=1}^{2}\pi_s exp\left(-\frac{(y - v_s)^2\rho^2}{2(v_2 - v_1)^2}\right)\right], \qquad (8)$$
$$\theta_1 = (v_1,\ v_2,\ \pi_1,\ \pi_2,\ \sigma^2),$$
$$\rho^2 = (v_2 - v_1)^2\sigma^{-2}.$$

Analytical expressions (7) and (8) differ only the constant multipliers before square brackets and not influencing the character and position of the extrema. Therefore the sufficient conditions theorems of the unimodality of the investigated mixture, proved for p=1 in [4] exist and for $p \geq 2$.

Theorem 1. The probability density (2) is unimodal, if $\rho \leq 2$.

Theorem 2. The probability density (2) for $\rho > 2$ and $\pi_1 \neq \pi_2$ is unimodal, if

$$|ln(\pi_1\pi_2^{-1})| \geq 2^{-1}\rho^2 + 2ln\left(2^{-1}\left(\rho + \sqrt{\rho^2 - 4}\right)\right).$$

Theorem 3. If the probability density (2) is bemodal for $\rho > 2$, $\pi_1 \neq \pi_2$ then the inequality

$$|ln(\pi_1\pi_2^{-1})| < 2^{-1}\rho^2 + 2ln\left(2^{-1}\left(\rho + \sqrt{\rho^2 - 4}\right)\right)$$

exists. At p=1, $\rho > 2$, $\pi_1 \neq \pi_2$ the necessary and the sufficient conditions of the unimodality and the bimodality for the investigated mixture are received in [5].

References

1. T. Anderson. An Introduction to Multivariate Statistical Analysis. New York, John Wiley & Sons, 1960.
2. N.N. Aprausheva. On the Condition of Unimodality and Bimodality of Two Normal Classes Mixture.

Reseaches on Probability-Statistical Modelling the Real Processes. Moskva, Tsentr. Economy-Mathematical Inst. Akad. Nauk SSSR, 1977 [in Russian].

3. N.N. Aprausheva, M.B. Radzhabova. Cotton-raw Classification by Statistical Algorithm. Moscow, Vychisl. Tsentr. Akad. Nauk. SSSR, 1990 [in Russian].
4. N.N. Aprausheva, S.V. Sorokin. On the Unimodality of a Simple Gaussian Mixture // J. Computational Mathematic and Mathematical Physics,. Vol. 44, No 5, 2004, pp. 785-793.
5. N. Aprausheva, S. Sorokin. On the unimodality and the bimodality of a Gaussian mixture of the two components 8th Intern. Conference on Pattern Recognition and Image Analysis. PRIA-8-2007. Yoshkar-Ola, the Russian Federation, 2007, pp. 14-16.
6. M. A. Carreira-Perpiñán, C. Williams. On the Number of Modes of a Gaussian Mixture. Inform. Res. Report EDI-INF-RR-0159. School of Inf. Univ. of Edinburg, 2003.
7. H. Cramér. Mathematical Methods of Statistics. Princeton University Press, Princeton, N.J., 1946.
8. Ya.A. Fomin, G.R. Tarlovski. Statistical Theory of Pattern Recognition. Moscow, “Radio i Svyaz”, 1986 [in Russian].