Generalized Gamma-Clustering from a Phase Transition Viewpoint: Connection between Gamma-modes and Gamma-means

Oita University of Nursing and Health Sciences, Akifumi Notsu
The Graduate University of Advanced Study, Katsuhiro Omae
The Institute of Mathematical Statistics, Shinto Eguchi

1. Introduction In [1] Gamma-modes and Gamma-means were proposed as robust clustering. Gamma-modes works well when true clusters are well separated, but otherwise it does not work well. On the other hand Gamma-means can find clusters even if true clusters are not well separated, however, Gamma-means suffers from unstability for initial point, similar to K-means. We discuss a connection between Gamma-means and Gamma-modes by a one-parameter τ, so that one parameter family of clustering procedures is derived.

2. Generalized Gamma-Clustering We employ a multivariate t-distribution

$$f_\nu(x-\mu) = c_\nu \Big\{ 1 + \frac{1}{\nu}\|x-\mu\|^2 \Big\}^{-(\nu+p)/2}, \tag{1}$$

for x of $\mathbb{R}^p$, where c_ν is the normalizing constant, and μ in $\mathbb{R}^p$ is the mean vector. Suppose $x_i \in \mathbb{R}^p, i = 1, 2, \ldots, n$. Consider a loss function

$$L(M_K, \tau, \gamma) = \sum_{i=1}^{n} \frac{1}{\tau} \log \left\{ \frac{1}{K} \sum_{k=1}^{K} \exp(-\tau f_\nu(x_i - \mu_k)^\gamma) \right\}, \tag{2}$$

where $M_K = (\mu_1, \ldots, \mu_K)$ and $\mu_k \in \mathbb{R}^p$. Then the cluster centers are defined by

$$\widehat{M}_K^{\tau,\gamma} = \operatorname*{argmin}_{M_K} L(M_K, \tau, \gamma).$$

Note that

$$\lim_{\tau \to -\infty} L(M_K, \tau, \gamma) = \sum_{i=1}^{n} \min_{1 \le k \le K} -f_\nu(x_i - \mu_k)^\gamma \tag{3}$$

and

$$\lim_{\tau \to 0} L(M_K, \tau, \gamma) = \frac{1}{K} \sum_{k=1}^{K} \sum_{i=1}^{n} -f_\nu(x_i - \mu_k)^\gamma. \tag{4}$$

The limiting loss function (3) leads to the Gamma-means; the limiting loss function (4) leads to the Gamma-modes. In this way, the one-parameter family of loss functions (2) includes both the Gamma-means and the Gamma-modes in the limiting sense. From a view point of phase transition we observe that a value of τ controls the hardness and separation of clustering [2]. A reasonable selection for the tuning parameter τ improves both weak points of Gamma-modes and Gamma-means.

References

[1] Akifumi Notsu and Shinto Eguchi. Robust clustering method in the presence of scattered observations. *Neural Computation*, 28(6):1141–1162, 2016.

[2] Kenneth Rose, Eitan Gurewitz, and Geoffrey C. Fox. Statistical mechanics and phase transitions in clustering. *Physical review letters*, 65(8), 1990.