Generalized Bhattacharyya and Chernoff upper bounds on Bayes error using quasi-arithmetic means

Pattern Recognition Letters 42: 25-34 (2014). BIBTEX:`J2014-GenBhatChernoff-PRL` [1]

In statistical hypothesis testing, we are given observations x that can emanate either from $X_1 \sim p_1$ or $X_2 \sim p_2$ with *a priori* weights w_1 and w_2, respectively. The *Bayes error* B_e for the cost design matrix $C = [c_{ij}]$ is related to the *total variation* metric distance $\mathrm{TV}(p,q) = \frac{1}{2}\int |p(x) - q(x)|\mathrm{d}x$ by $B_e = \frac{a_1+a_2}{2} - \mathrm{TV}(a_1p_1, a_2p_2)$ with $a_1 = w_1(c_{11} + c_{21})$ and $a_2 = w_2(c_{12} + c_{22})$. The identity simplifies for probability of error P_e to $P_e = \frac{1}{2} - \mathrm{TV}(w_1p_1, w_2p_2)$. For multivariate normals $X_1 \sim N(\mu_1, \Sigma)$ and $X_2 \sim N(\mu_2, \Sigma)$, we have $P_e = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{1}{2\sqrt{2}}\|(\Sigma^{-1})^{\frac{1}{2}}(\mu_2 - \mu_1)\|\right)$. Usually, it is difficult to have closed form formula for the Bayes' error or the probability of error so that upper bounds are considered. We define novel affinity coefficients and divergences that play a role on upper bounding the Bayes' error using quasi-arithmetic means. The quasi-arithmetic weighted mean $M_f(a,b;\alpha) = f^{-1}(\alpha f(a) + (1-\alpha)f(b))$ of two real values a and b for a strictly monotonic function f satisfies the interness property: $\min(a,b) \leq M_f(a,b;\alpha) \leq \max(a,b)$. The *Chernoff-type similarity coefficient* (affinity) for a strictly monotonous function f is defined by $\rho_*^f(p_1,p_2) = \min_{\alpha\in[0,1]}\int M_f(p_1(x),p_2(x);\alpha)\mathrm{d}x \leq 1$ and we define the *generalized Chernoff information* by $C_f(p_1,p_2) = -\log\rho_*^f(p_1,p_2) = \max_{\alpha\in[0,1]} -\log\int M_f(p_1(x),p_2(x);\alpha)\mathrm{d}x \geq 0$. The *generalized skew Bhattacharyya-type similarity coefficient* (affinity) is $\rho_\alpha^f(p_1,p_2) = \int M_f(p_1(x),p_2(x);\alpha)\mathrm{d}x \leq 1$, and the *generalized skew Bhattacharyya-type divergence* is $B_\alpha^f = -\log\rho_\alpha^f(p_1,p_2)$. The *generalized Bhattacharyya coefficient* $\rho^f(p_1,p_2) = \int M_f(p_1(x),p_2(x);\frac{1}{2})\mathrm{d}x$ and divergence $B^f(p_1,p_2) = -\log\rho^f(p_1,p_2)$. Those definitions generalize the traditional cases where $f(x) = \log x$. The Chernoff bounds are interpreted as the minimization of a geometric weighted mean, and a generic Chernoff bound construction is expressed using quasi-arithmetic weighted means. Using quasi-arithmetic means, we can bound the probability of error as $P_e = \int\min(w_1p_1(x), w_2p_2(x))\mathrm{d}x \leq \int M_f(w_1p_1(x), w_2p_2(x);\alpha)\mathrm{d}x$. The upper bound proves useful for well-chosen f yielding closed-form expressions. We illustrate this upper bounding technique with (1) geometric means and the Chernoff bound for exponential families, (2) Harmonic means and upper bounds for Cauchy distributions, and (3) Pearson type VII distributions and central multivariate t -distributions with power means. Experiments assess those bounds (codes in R).

References

[1] Frank Nielsen. Generalized Bhattacharyya and Chernoff upper bounds on Bayes error using quasi-arithmetic means. *Pattern Recognition Letters*, 42(0):25 – 34, 2014.