# A measure of discrimination and its geometric properties

Rajesh Sundaresan[1]
Qualcomm Inc.
3480 Granada Avenue 270,
Santa Clara, CA 95051, USA
e-mail: rajeshs@qualcomm.com

*Abstract* — **We study a measure of discrimination that arises in the context of redundancy in guessing the realization of a random variable. This discrimination measure satisfies a Pythagorean-type inequality. The analog of this inequality for Kullback-Leibler divergence is well-known.**

Let $P$ and $Q$ be two probability mass functions (PMFs) on a finite set $\mathcal{X}$. The measure of discrimination that we consider here is $L_\alpha(P,Q)$ defined as

$$\frac{\alpha}{1-\alpha}\log\left(\sum_{x\in\mathcal{X}}P(x)\left[\sum_{a\in\mathcal{X}}\left(\frac{Q(a)}{Q(x)}\right)^{\alpha}\right]^{\frac{1-\alpha}{\alpha}}\right)-H_\alpha(P), \quad (1)$$

where $\alpha>0$, and $H_\alpha(P)$ is the Rényi entropy of order $\alpha>0$. For $\alpha\in(0,1)$, $L_\alpha(P,Q)$ can be interpreted as follows.

Let $X$ be a random variable on the finite set $\mathcal{X}$ with probability mass function (PMF) given by $(P(x): x\in\mathcal{X})$. We wish to guess the realization of this random variable $X$ by asking questions of the form "Is $X$ equal to $x$?", stepping through the elements of $\mathcal{X}$, until the answer is "Yes" ([1], [2]). For a given guessing strategy $G$, let $G(x)$ denote the number of guesses required when $X=x$. The strategy that minimizes the expected number of guesses $E[G(X)]$, or more generally $||G(X)||_\rho=(E[G(X)^\rho])^{1/\rho}$ for $\rho>0$, proceeds in the decreasing order of source probabilities.

Let $\alpha=1/(1+\rho)$. The minimum value of $\log||G(X)||_\rho$ is within $\log(1+\ln|\mathcal{X}|)$ of the second term in (1), i.e., $H_\alpha(P)$. If however guessing is done according to some suboptimal guessing function $G$, the value of $\log||G(X)||_\rho$ lies within $\log(1+\ln|\mathcal{X}|)$ of the first term in (1), where $Q$ is the PMF such that $Q(\cdot)\propto G(\cdot)^{-(1+\rho)}$.

$L_\alpha(P,Q)$ is therefore a measure of the *redundancy* in $\log||G(X)||_\rho$ as a result of a PMF mismatch. $L_\alpha(P,Q)\geq 0$ and equals 0 if and only if $P=Q$. $L_\alpha(P,Q)$ does not satisfy many interesting properties that hold for the Kullback-Leibler divergence; for e.g., $L_\alpha(P,Q)$ does not satisfy the data processing inequality. However, $L_\alpha(P,Q)$ and the Kullback-Leibler divergence do share a Pythagorean-type inequality property. To see this, we proceed along the lines of [3].

Given a PMF $R$ on the finite set $\mathcal{X}$, the set of PMFs on $\mathcal{X}$

$$B(R,r)\triangleq\{P: L_\alpha(P,R)<r\}, \quad 0<r\leq\infty,$$

is called an $L_\alpha$-sphere (or ball) with center $R$ and radius $r$. If $\mathcal{E}$ is a closed and convex set of PMFs on $\mathcal{X}$ intersecting $B(R,\infty)$, a PMF $Q\in\mathcal{E}$ satisfying

$$L_\alpha(Q,R)=\min_{P\in\mathcal{E}}L_\alpha(P,R)$$

exists and is called the $L_\alpha$-projection of $R$ on $\mathcal{E}$.

The following are generalizations of [3, Lemma 2.1, Theorem 2.2] where $L_\alpha(P,Q)$ plays the role of squared Euclidean distance (analogous to the Kullback-Leibler divergence).

**Theorem 1** *Let $\alpha\in(0,1)\cup(1,\infty)$ and $\mathcal{X}$ a finite set.*

*(a) If $L_\alpha(P,Q)$ and $L_\alpha(Q,R)$ are finite, the segment joining $P$ and $Q$ does not intersect the $L_\alpha$-sphere $B(R,r)$ with radius $r=L_\alpha(Q,R)$, i.e.,*

$$L_\alpha(P_\lambda,R)\geq L_\alpha(Q,R),$$

*for each*

$$P_\lambda=\lambda P+(1-\lambda)Q, \quad 0\leq\lambda\leq 1,$$

*if and only if*

$$L_\alpha(P,R)\geq L_\alpha(P,Q)+L_\alpha(Q,R).$$

*(b)* (Tangent hyperplane) *If $L_\alpha(Q,R)$ and $L_\alpha(P,R)$ are finite and*

$$Q=\lambda P+(1-\lambda)S, \quad 0<\lambda<1, \quad (2)$$

*the segment joining $P$ and $S$ does not intersect $B(R,r)$ (with $r=L_\alpha(Q,R)$) if and only if*

$$L_\alpha(P,R)=L_\alpha(P,Q)+L_\alpha(Q,R).$$

We call $Q$ an algebraic innerpoint of $\mathcal{E}$ if for every $P\in\mathcal{E}$ there exist $\lambda$ and $S$ satisfying (2). Theorem 1 then leads to the following result on $L_\alpha$-projection.

**Theorem 2** (Projection Theorem) *Let $\alpha\in(0,1)\cup(1,\infty)$, $\mathcal{X}$ a finite set, and $\mathcal{E}$ a closed and convex set of PMFs on $\mathcal{X}$. Let $R\notin\mathcal{E}$. A PMF $Q\in\mathcal{E}\cap B(R,\infty)$ is the $L_\alpha$-projection of $R$ on the convex set $\mathcal{E}$ if and only if every $P\in\mathcal{E}\cap B(R,\infty)$ satisfies*

$$L_\alpha(P,R)\geq L_\alpha(P,Q)+L_\alpha(Q,R). \quad (3)$$

*If the $L_\alpha$-projection $Q$ is an algebraic inner point of $\mathcal{E}$, then every $P\in\mathcal{E}\cap B(R,\infty)$ satisfies (3) with equality.*

The uniqueness of the $L_\alpha$-projection on $\mathcal{E}$ follows straightforwardly from Theorem 2 and the fact that $L_\alpha(Q_1,Q_2)=0$ implies $Q_1=Q_2$.

## References

[1] J.L.Massey, "Guessing and entropy," in *Proc. 1994 IEEE Int. Symp. on Information Theory* (Trondheim, Norway, 1994), p. 204.

[2] E.Arikan, "An inequality on guessing and its application to sequential decoding," *IEEE Trans. Inform. Theory*, vol. IT-42, pp. 99-105, Jan. 1996.

[3] I.Csiszar, "I-divergence geometry of probability distributions and minimization problems," *The Annals of Probability*, vol. 3, pp. 146-158, 1975.

[1]This work was supported in part by the National Science Foundation under Grant NCR-9523805 002



0-7803-7501-7/02/$17.00©2002 IEEE.