HYBRID MUTUAL INFORMATION BETWEEN DIRECTIONAL AND LINEAR VARIABLES

CIG

IGNACIO LEGUEY, SHOGO KATO, CONCHA BIELZA AND PEDRO LARRAÑAGA COMPUTATIONAL INTELLIGENCE GROUP

POLITÉCNICA "Engineering the future" INTERNATIONAL CAMPUS OF EXCELLENCE

MOTIVATION

- **Mutual information** [1]: general measure that determines the similarity between the joint distribution and the product of their marginal distributions.
- **Circular mutual information** [2]: mutual information in directional domains.
  - Numerical methods to approach it.
  - Suitable when the variables follow wrapped Cauchy distributions [3].

  –> *Improvement*
- **Directional-linear** data is ubiquitous in science –> *directional-linear mutual information*

JOHNSON & WEHRLY BIVARIATE DISTRIBUTIONS [4][5]

- Let Θ be a specific directional variable and let X be a specific linear variable, with $f(\theta)$ and $f(x)$ their marginal density functions and $F(\theta)$ and $F(x)$ their cumulative density functions. Then, the joint density function can be expressed as [4]

$$f(\theta, x) = 2\pi\delta[2\pi F(\theta) \pm 2\pi F(x)]f(\theta)f(x).$$

  where $0 \leq \theta < 2\pi$, $-\infty < x < \infty$ and $\delta(.)$ is a density on the circle.

- Likewise, let Ψ a specific directional variable with $f(\psi)$ its marginal density function and $F(\psi)$ its CDF. Then, the joint density function of Θ and Ψ can be expressed as [5]

$$f(\theta, \psi) = 2\pi\delta[2\pi F(\theta) \pm 2\pi F(\psi)]f(\theta)f(\psi).$$

  where $0 \leq \theta, \psi < 2\pi$.

DENSITY ON THE CIRCLE: WRAPPED CAUCHY [6]

A directional random variable Θ, that has probability density function

$$g(\theta) = \frac{1}{2\pi}\left[1 + \frac{2\varepsilon^2\{\varepsilon\cos(\theta-\mu)-\alpha\}}{\varepsilon^2+\alpha^2+\beta^2-2\varepsilon\{\alpha\cos(\theta-\mu)+\beta\sin(\theta-\mu)\}}\right],$$

where μ is the location parameter, ε is the concentration parameter, α is the circular kurtosis and β is the circular skewness. When $\alpha = \varepsilon^2$ and $\beta = 0$, $g(\theta)$ follows wrapped Cauchy distribution $wC(\mu, \varepsilon)$ with density

$$g(\theta) = \frac{1}{2\pi}\frac{1-\varepsilon^2}{1+\varepsilon^2-2\varepsilon\cos(\theta-\mu)}, \quad \theta, \mu \in (-\pi, \pi], \varepsilon \in [0, 1)$$

When $\varepsilon = 0$, $g(\theta)$ follows the circular uniform distribution, otherwise is unimodal and symmetric about μ.

FUTURE WORK

- Extending these measures to the multivariate case in order to understand the interactions among many random variables regardless of their nature.
- Develop Bayesian networks and supervised classification models using this measures.

MAIN RESULTS

- Let $f(\theta, x)$ be the joint density function defined by Johnson & Wehrly [4], with $\delta(.)$ as a wrapped Cauchy presented in Kato & Jones [6]. Then, the hybrid mutual information (HMI) between a directional specific variable Θ and a specific linear variable X is

$$\text{HMI}(\Theta, X) = -\log(1 - \varepsilon_\delta^2).$$

  where ε_δ is the concentration parameter from $\delta(.)$, which is defined as

$$\varepsilon_\delta = \frac{1}{n}\left(|\sum_{j=1}^{n} e^{i(2\pi F(\theta_j)-(2\pi F(x_j))}| - |\sum_{j=1}^{n} e^{i((2\pi F(\theta_j)+(2\pi F(x_j))}|\right).$$

  with $F(\theta)$ and $F(x)$ the CDFs of Θ and X respectively.

- Likewise, let $f(\theta, \psi)$ be the joint density function for directional variables defined by Johnson & Wehrly [5]. Then, the circular mutual information (CMI) between two specific directional variables Θ and Ψ is

$$\text{CMI}(\Theta, \Psi) = -\log(1 - \varepsilon_\delta^2).$$

  where ε_δ is the concentration parameter of $\delta(.)$.

- These two measures share two important properties:
  - Both are expressed in a closed form, therefore are computationally very fast and easy to handle.
  - Allows the use of variables that follows any directional or linear distribution to compute them.

REFERENCES

[1] J.A. THOMAS AND T.M COVER, Elements of information theory, *John Wiley & Sons* (1991).

[2] I. LEGUEY, C. BIELZA AND P. LARRAÑAGA, Tree-structured Bayesian networks for wrapped Cauchy directional distributions. *Lecture Notes in Computer Science*. Springer, **9868** (2016) 207–216.

[3] S. KATO AND A. PEWSEY, A Möbius transformation-induced distribution on the torus, *Biometrika* **102(2)** (2015), 359–370.

[4] R.A. JOHNSON AND T.E. WEHRLY, Some angular-linear distributions and related regression models, *Journal of the American Statistical Association* **73(363)** (1978), 602–606.

[5] T.E. WEHRLY AND R.A. JOHNSON, Bivariate models for dependence of angular observations and a related Markov process, *Biometrika* **67(1)** (1980), 255.

[6] S. KATO AND M.C. JONES, A tractable and interpretable four-parameter family of unimodal distributions on the circle, *Biometrika* **102(1)** (2015), 181.

CONTACT DETAILS

If you are interested in further information, please do not hesitate to contact me by e-mail: **ig.leguey@upm.es**