

## 2ª Reunión: Preguntas

- ¿Los datos en la polysphera de sunspots-bird, por ejemplo, podrían ser los diferentes nacimientos de sunspots para cada ciclo solar? Me cuesta ver como tener datos en la polyesfera. En Lunga2013, el ejemplo se basa en hyperspectral images, teniendo para un mismo píxel diferentes valores. Aquí si que lo visualizo mejor, sin embargo, para las coordenadas en la esfera me cuesta verlo.
- En nuestro caso, ¿Tanto la high-dimension como la low-dimension se modelan con una distribución de Cauchy en la (poly-)esfera?
- ¿Qué distancia deberíamos usar para modelar la probabilidad de que una observación  $i$  escoja a  $j$  como vecino? He leído que la distancia Euclídea no funciona muy bien para nonlinear manifolds. En la literatura, he leído que se han tenido en cuenta a veces la distancia geodésica, la distancia del coseno.
- He desarrollado el cálculo de  $P$ :

$$p_{j|i} = \frac{p_{ji}}{p_i} = \frac{\frac{\Gamma\{(d+1)/2\}}{2\pi^{(d+1)/2}} \left( \frac{1-\rho^2}{1+\rho^2-2\rho\boldsymbol{\mu}'S_c(\mathbf{x}_i, \mathbf{x}_j)} \right)^d}{\frac{\Gamma\{(d+1)/2\}}{2\pi^{(d+1)/2}} \sum_{k \neq i} \left( \frac{1-\rho^2}{1+\rho^2-2\rho\boldsymbol{\mu}'S_c(\mathbf{x}_i, \mathbf{x}_k)} \right)^d}$$

arity is denoted by  $S_c(\mathbf{x}_i, \mathbf{x}_j) = \cos(\theta) = \frac{\mathbf{x}_i \mathbf{x}_j}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|}$ .

Me faltaría tener en cuenta la perplexity a la hora de calcular los parámetros de la función de densidad de la distribución Cauchy en la esfera. Para que la perplexity fuera la misma para todas las observaciones. ¿Qué literatura podría leer más sobre esta distribución? Tan solo he encontrado la de la pasada reunión que también comentaba sobre Möbius transformation.

- ¿Existen funciones ya implementadas (de la distribución de Cauchy en la esfera) en R de la pdf, la cdf, del quantil y de la generación de datos de esta distribución? De no ser así, ¿Estaría bien desarrollarlas? Tan solo he encontrado los paquetes Directional (función `spcauchy.mle` para estimar los parámetros mediante MLE) y el paquete `DirStats` para trabajar con datos direccionales.
- En Lunga 2013, se habla de preservar la Hellinger distance y del coeficiente de Bhattacharyya. ¿Debería investigar más acerca de estos términos?
- En Kato 2009 se habla de la distribución de un par de vectores unitarios generados por una Brownian motion, la única relación que existe es con la sSNE que utiliza la distribución Exit para modelar la low-dimension no?
- Formato del documento, tabla de contenidos. ¿Podría trabajar ya en algo teórico, por ejemplo, state of the art?