Non-linear Embeddings in Hilbert Simplex Geometry

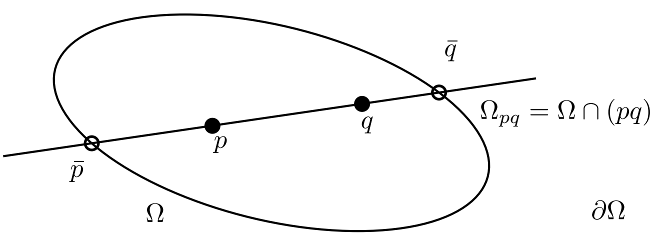
Frank Nielsen
Sony Computer Science Laboratories
Tokyo, Japan

Ke Sun
CSIRO Data61
Sydney, Australia

Contributions:

- Simple proof of monotonicity of Hilbert distance
- Connection of *Hilbert distance* with *Aitchison distance*
- Differentiable approximation of Hilbert distance
- Application to non-linear embedding: experimentally fast, robust, and competitive

Open bounded convex Ω of $\mathbb{R}^d$:



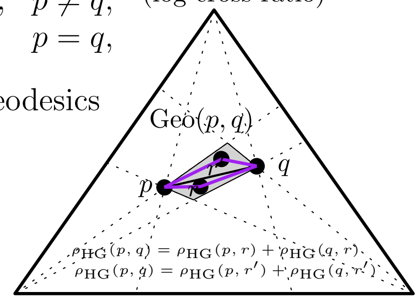
$\rho^{\Omega}_{\mathrm{HG}}(p,q) = \rho^{\Omega_{pq}}_{\mathrm{HG}}(p,q)$

Hilbert metric distance: Symmetrize Funk distance

$$\rho^{\Omega}_{\mathrm{HG}}(p,q) := \rho^{\Omega}_{\mathrm{FD}}(p,q) + \rho^{\Omega}_{\mathrm{FD}}(q,p) = \begin{cases} \log \frac{\|p-\bar{q}\|\,\|q-\bar{p}\|}{\|p-\bar{p}\|\,\|q-\bar{q}\|}, & p \neq q, \\ 0 & p = q. \end{cases}$$

$$\rho^{\Omega}_{\mathrm{HG}}(p,q) = \begin{cases} \log \mathrm{CR}(\bar{p},p;q,\bar{q}), & p \neq q, \\ 0 & p = q, \end{cases}$$ (log cross-ratio)

Straight line segments = geodesics but geodesics not unique:

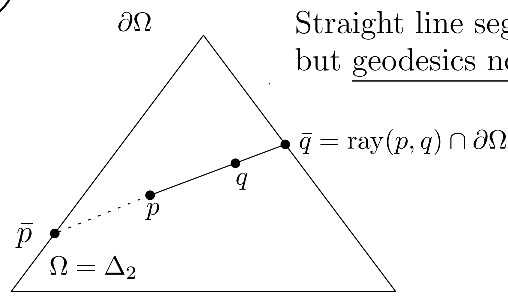


Hilbert simplex distance:

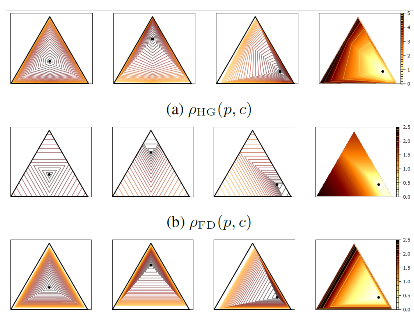
$$\Delta_d := \left\{(x_1,\ldots,x_d) \in \mathbb{R}^d_{++} \ :\ \sum_{i=1}^d x_i = 1\right\}$$

$$\rho_{\mathrm{FD}}(p,q) = \log \max_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}$$

$$\rho_{\mathrm{HG}}(p,q) = \log \frac{\max_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}{\min_{i\in\{1,\ldots,d\}} \frac{p_i}{q_i}}$$



$$F_\Omega(p:q) = \log \frac{\|p-\bar{q}\|}{\|q-\bar{q}\|} = \log \max_i \frac{p_i}{q_i}$$


(a) $\rho_{\mathrm{HG}}(p,c)$
(b) $\rho_{\mathrm{FD}}(p,c)$
(c) $\rho_{\mathrm{FD}}(c,p)$

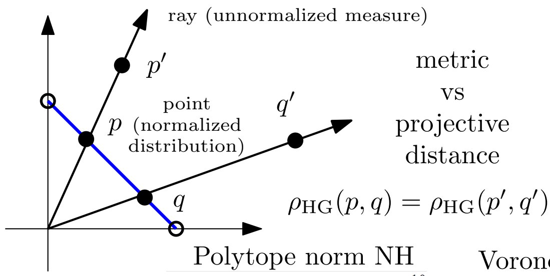
Aitchison distance:

$$\rho_{\mathrm{Aitchison}}(p,q) := \sqrt{\sum_{i=1}^d \left(\log \frac{p_i}{G(p)} - \log \frac{q_i}{G(q)}\right)^2}$$

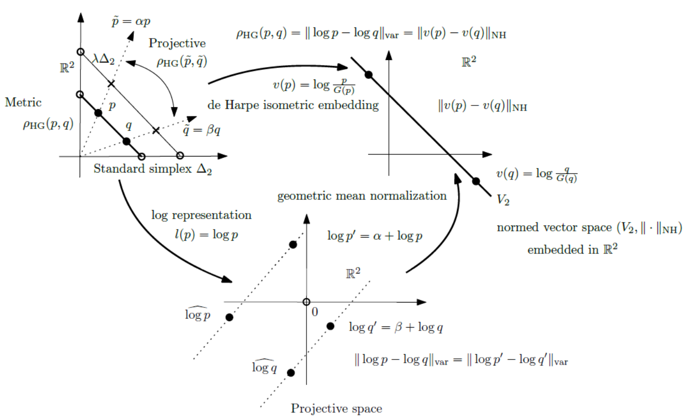
geometric mean:

$$G(p) = \left(\prod_{i=1}^d p_i\right)^{\frac{1}{d}} = \exp\left(\frac{1}{d}\sum_{i=1}^d \log p_i\right)$$

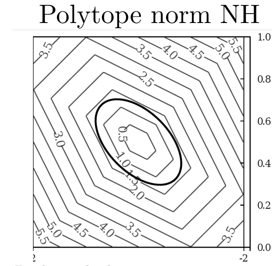
Positive orthant cone $\mathbb{R}^2$



$\rho_{\mathrm{HG}}(p,q) = \rho_{\mathrm{HG}}(p',q')$

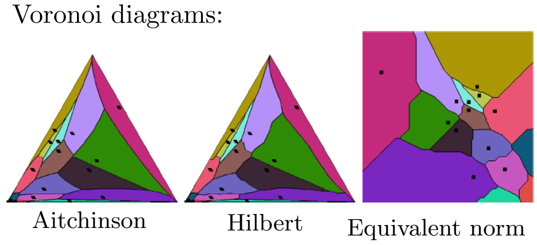


Polytope norm NH



$V^3 = \{v \in \mathbb{R}^3 \ :\ \sum_{i=1}^3 v^i = 0\}$ of $\mathbb{R}^3$

Voronoi diagrams:



Differentiable approximation:

$$\tilde{\rho}_{\mathrm{LSE}^T}(p,q) = \frac{1}{T}\log\left(\sum_i \left(\frac{p_i}{q_i}\right)^T\right)\left(\sum_i \left(\frac{q_i}{p_i}\right)^T\right)$$

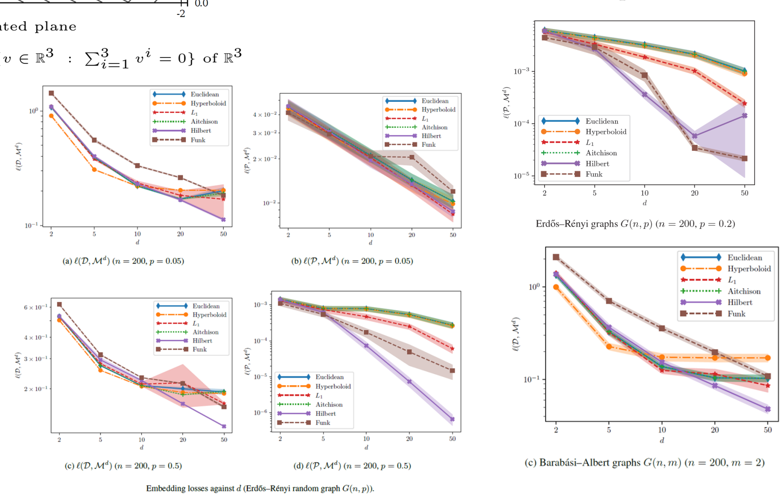
$$\lim_{T\to\infty} \tilde{\rho}^{\mathrm{LSE}^T}(p,q) = \rho(p,q)$$

Loss functions:
Empirical average KLD
+ Adam optimizer

$$\ell(\mathcal{D}, \mathcal{M}^d) := \inf_{Y \in (\mathcal{M}^d)^n} \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n (\mathcal{D}_{ij} - \rho_{\mathcal{M}}(y_i, y_j))^2$$

$$\ell(\mathcal{P}, \mathcal{M}^d) := \inf_{Y \in (\mathcal{M}^d)^n} \frac{1}{n}\sum_{i=1}^n \sum_{j:j\neq i} \mathcal{P}_{ij} \log \frac{\mathcal{P}_{ij}}{q_{ij}(Y)},$$

$$q_{ij}(Y) := \frac{\exp(-\rho^2_{\mathcal{M}}(y_i,y_j))}{\sum_{j:j\neq i}\exp(-\rho^2_{\mathcal{M}}(y_i,y_j))},$$



(a) $\ell(\mathcal{D}, \mathcal{M}^d)$ ($n = 200$, $p = 0.05$)
(b) $\ell(\mathcal{P}, \mathcal{M}^d)$ ($n = 200$, $p = 0.05$)
(c) $\ell(\mathcal{D}, \mathcal{M}^d)$ ($n = 200$, $p = 0.5$)
(d) $\ell(\mathcal{P}, \mathcal{M}^d)$ ($n = 200$, $p = 0.5$)
Embedding losses against d (Erdős–Rényi random graph $G(n,p)$).
Erdős–Rényi graphs $G(n,p)$ ($n = 200$, $p = 0.2$)
(c) Barabási–Albert graphs $G(n,m)$ ($n = 200$, $m = 2$)

See experiments in arxiv:2203.11434