

Taxonomy of principal distances

Euclidean geometry

Euclidean distance
 $d_2(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_i (p_i - q_i)^2}$ (Pythagoras' theorem circa 500 BC)



Hamming distance
 $(|\{i : p_i \neq q_i\}|)$

Manhattan distance
 $d_1(\mathbf{p}, \mathbf{q}) = \sum_i |p_i - q_i|$ (city block-taxi cab)



Statistical geometry

Physics entropy JK^{-1}
 $-k \int p \log p$ (Boltzmann-Gibbs 1878)



Additive entropy

cross-entropy
 conditional entropy
 mutual information
 (chain rules)

Information entropy
 $H(p) = -\int p \log p$ (C. Shannon 1948)



Minkowski distance (L_k -norm)
 $d_k(\mathbf{p}, \mathbf{q}) = \sqrt[k]{\sum_i |p_i - q_i|^k}$
 (H. Minkowski 1864-1909)
Space-time geometry



Mahalanobis metric (1936)
 $d_\Sigma = \sqrt{(\mathbf{p} - \mathbf{q})^T \Sigma^{-1} (\mathbf{p} - \mathbf{q})}$



Quadratic distance
 $d_Q = \sqrt{(\mathbf{p} - \mathbf{q})^T \mathbf{Q} (\mathbf{p} - \mathbf{q})}$



Riemannian metric tensor
 $\int \sqrt{g_{ij} \frac{dx_i}{ds} \frac{dx_j}{ds}} ds$
 (B. Riemann 1826-1866.)

Non-Euclidean geometries

Fisher information (local entropy)
 $\mathbf{I}(\theta) = E[\left(\frac{\partial}{\partial \theta} \ln p(X|\theta)\right)^2]$
 (R. A. Fisher 1890-1962)



I -projection

Kullback-Leibler divergence
 $KL(\mathbf{p}||\mathbf{q}) = \int p \log \frac{p}{q} = E_p[\log \frac{p}{q}]$
 (relative entropy, 1951)



$H(p) = KL(p||u)$



Life negative entropy



Jeffrey divergence
 (Jensen-Shannon)

Bhattacharya distance (1967)
 $d(p, q) = -\ln \sqrt{\int (\sqrt{p} - \sqrt{q})^2}$



Kolmogorov
 $K(p||q) = \int |q - p|$
 (Kolmogorov-Smirnoff max $|p - q|$)

Matsushita distance (1956)
 $M_\alpha(p, q) = \sqrt[\alpha]{\int |q^\alpha - p^\alpha|}$



exponential families

Chernoff divergence (1952)
 $C_\alpha(p||q) = -\ln \int p^\alpha q^{1-\alpha}$



Itakura-Saito divergence
 $IS(\mathbf{p}||\mathbf{q}) = \sum_i (\frac{p_i}{q_i} - \log \frac{p_i}{q_i} - 1)$
 (Burg entropy)



Rényi divergence (1961)
 $H_\alpha = \frac{1}{\alpha(1-\alpha)} \log \int f^\alpha$
 $R_\alpha(\mathbf{p}||\mathbf{q}) = \frac{1}{\alpha(\alpha-1)} \ln \int p^\alpha q^{1-\alpha}$
 (additive entropy)



Hellinger
 $H(p||q) = \sqrt{\int (\sqrt{p} - \sqrt{q})^2}$
 $= \sqrt{2(1 - \int \sqrt{pq})}$



χ^2 test
 $\chi^2(p||q) = \int \frac{(q-p)^2}{p}$
 (K. Pearson, 1857-1936)



Neyman

Csiszár' f -divergence
 $D_f(p||q) = \int p f(\frac{q}{p})$
 (Ali& Silvey 1966, Csiszár 1967)

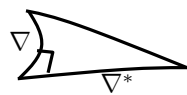


Dual div. *-conjugate ($f^*(y) = yf(1/y)$)
 $D_{f^*}(p||q) = D_f(q||p)$



Information geometries

Amari α -divergence (1985)
 $f_\alpha(x) = \begin{cases} x \log x & \alpha = 0 \\ -\log x & \alpha = 1 \\ \frac{4}{1-\alpha^2} (1-x \frac{1+\alpha}{2}) & -1 < \alpha < 1 \end{cases}$



Quantum geometry

Quantum entropy
 $S(\rho) = -k \text{Tr}(\rho \log \rho)$
 (Von Neumann 1927)



Von Neumann divergence
 $D(\mathbf{P}||\mathbf{Q}) = \text{Tr}(\mathbf{P}(\log \mathbf{P} - \log \mathbf{Q}) - \mathbf{P} + \mathbf{Q})$

Burbea-Rao
 (incl. Jensen-Shannon)
 $J_F(p, q) = \frac{f(p)+f(q)}{2} - f\left(\frac{p+q}{2}\right)$

Log Det divergence
 $D(\mathbf{P}||\mathbf{Q}) = \langle \mathbf{P}, \mathbf{Q}^{-1} \rangle - \log \det \mathbf{P} \mathbf{Q}^{-1} - \dim \mathbf{P}$

Permissible Bregman divergences
 (Nock & Nielsen, 2007)

Non-additive entropy

Tsallis entropy (1998)
 (Non-additive entropy)
 $T_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} (\sum_i p_i^\alpha - 1)$
 $T_\alpha(p||q) = \frac{1}{1-\alpha} (1 - \int \frac{p^\alpha}{q^{\alpha-1}})$



Earth mover distance
 (EMD 1998)

Algorithmic geometry?

Distance between two algorithms ?

Kolmogorov complexity