Visualizing Bregman Voronoi diagrams*

Frank Nielsen
Sony Computer Science Labs
FRL
Tokyo, Japan

Jean-Daniel Boissonnat
INRIA
GEOMETRICA
Sophia-Antipolis, France

Richard Nock
Université Antilles-Guyane
CEREGMIA
Martinique, France

ABSTRACT

Voronoi diagrams are fundamental geometric structures that partition the space into elementary regions of influence defining discrete proximity graphs and dually well-shaped Delaunay triangulations [1]. In this video, we explain and illustrate a recent generalization of Voronoi diagrams [3] to a wide class of distortion measures called Bregman divergences [2].

Categories and Subject Descriptors

I.3.5 [**Computer Graphics**]: Computational Geometry and Object Modeling—*Geometric algorithms, languages, and systems*

General Terms

Algorithms, Theory

Keywords

Computational information geometry, Voronoi diagrams, Bregman divergences

1. BREGMAN DIVERGENCES

For any two points $\mathbf{p}$ and $\mathbf{q}$ of $\mathcal{X} \subseteq \mathbb{R}^d$, the Bregman divergence $D_F(\cdot||\cdot) : \mathbb{R}^d \mapsto \mathbb{R}$ of $\mathbf{p}$ to $\mathbf{q}$ associated to a strictly convex and differentiable function $F : \mathbb{R}^d \mapsto \mathbb{R}$ (called the *generator function* of the divergence) is defined as $D_F(\mathbf{p}||\mathbf{q}) \stackrel{\text{def}}{=} F(\mathbf{p}) - F(\mathbf{q}) - \langle \mathbf{\nabla} F(\mathbf{q}), \mathbf{p} - \mathbf{q} \rangle$, where $\mathbf{\nabla} F = [\frac{\partial F}{\mathrm{d}x_1} \ ... \ \frac{\partial F}{\mathrm{d}x_d}]^T$ denotes the gradient operator, and $\langle \cdot, \cdot \rangle$ the inner product. Informally speaking, Bregman divergence D_F is the *tail* of the Taylor expansion of F (Fig. 1). The Bregman divergence $D_F(\mathbf{p}||\mathbf{q})$ is geometrically measured as the vertical distance between $F(\mathbf{p})$ and the hyperplane $H_{\mathbf{q}}$ tangent to $\mathcal{F} : z = F(\mathbf{x})$ at point $\mathbf{q}$: $D_F(\mathbf{p}||\mathbf{q}) = F(\mathbf{p}) - H_{\mathbf{q}}(\mathbf{p})$. Bregman divergences are not necessarily symmetric nor do they satisfy the triangle inequality.

*`www.csl.sony.co.jp/person/nielsen/BregmanVoronoi/`

SoCG June 6-8, 2007, Gyeongju, South Korea

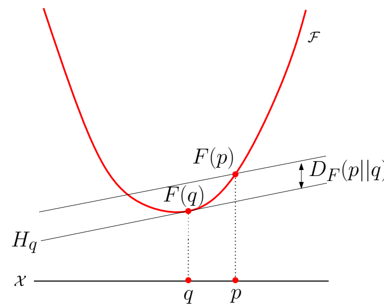

Figure 1: Visualizing Bregman divergences.

Bregman divergences admit dual Bregman divergences induced by convex conjugation. The Legendre transformation makes use of the duality relationship between points and lines to associate to F a convex conjugate function $F^* : \mathbb{R}^d \mapsto \mathbb{R}$ given by $F^*(\mathbf{y}) = \sup_{\mathbf{x} \in \mathcal{X}} \{\langle \mathbf{y}, \mathbf{x} \rangle - F(\mathbf{x})\}$. The supremum is reached at the *unique* point where the gradient of $G(\mathbf{x}) = \langle \mathbf{y}, \mathbf{x} \rangle - F(\mathbf{x})$ vanishes or, equivalently, when $\mathbf{y} = \mathbf{\nabla} F(\mathbf{x})$.

2. ELEMENTS OF BREGMAN GEOMETRY

Because Bregman divergences may not be symmetric, we first define two types of Bregman balls (and bounding spheres) as: $B_F(\mathbf{c}, r) = \{\mathbf{x} \in \mathcal{X} \mid D_F(\mathbf{x}||\mathbf{c}) \leq r\}$ (first-type), and $B'_F(\mathbf{c}, r) = \{\mathbf{x} \in \mathcal{X} \mid D_F(\mathbf{c}||\mathbf{x}) \leq r\}$ (second-type).

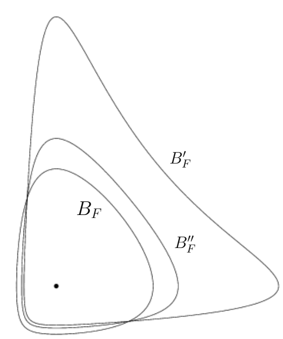

Figure 2: Three-types of Bregman spheres for the Itakura-Saito divergence (generator: Burg entropy $F(\mathbf{x}) = -\sum_i \log x_i$).

Second-type balls may not be convex although first-type balls are *always* convex. It is also convenient to symmetrize

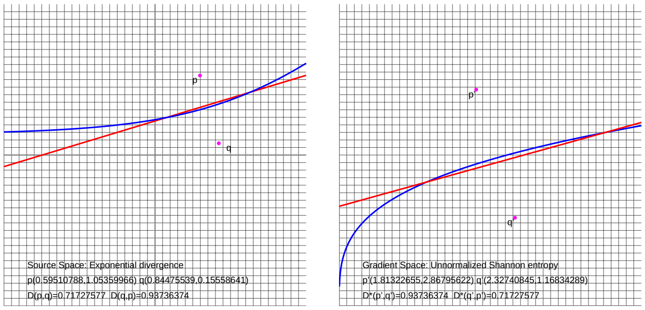

Figure 3: Bregman bisectors: first-type linear bisector (red) and second-type curved bisector (blue) are displayed for the exponential loss/unnormalized Shannon entropy (zoom in to read text please).

the Bregman divergence $S_F(\mathbf{p},\mathbf{q}) = \frac{1}{2}(D_F(\mathbf{p}||\mathbf{q}) + D_F(\mathbf{q}||\mathbf{p}))$, and define a third-type ball $B''_F = \{\mathbf{x} \in \mathcal{X} \mid S_F(\mathbf{c},\mathbf{x}) \leq r\}$ (Figure 2). It has been shown that the third-type symmetrized Bregman divergence is a "particular" divergence in higher dimensions [3]. Thus we concentrate on the first two type structures and define accordingly the Bregman bisectors: $H_F(\mathbf{p},\mathbf{q}) = \{\mathbf{x} \in \mathcal{X} \mid D_F(\mathbf{x}||\mathbf{p}) = D_F(\mathbf{x}||\mathbf{q})\}$ and $H'_F(\mathbf{p},\mathbf{q}) = \{\mathbf{x} \in \mathcal{X} \mid D_F(\mathbf{p}||\mathbf{x}) = D_F(\mathbf{q}||\mathbf{x})\}$. The first-type Bregman bisector is always an hyperplane. The second-type Bregman bisector, although curved in the primal space, is linear in the dual gradient space, as displayed in Figure 3.

3. COMPUTING DIAGRAMS

Figure 4 shows a Bregman Voronoi diagram for the Itakura-Saito divergence. The first-type Bregman Voronoi diagram is affine and can be computed from the minimization diagrams of corresponding set of functions $D_i(\mathbf{x}) = D_F(\mathbf{x}||\mathbf{p}_i)$. Each function D_i being an hyperplane of equation:
$-\langle \mathbf{x}, \nabla F(\mathbf{p}_i)\rangle - F(\mathbf{p}_i) + \langle \mathbf{p}_i, \nabla F(\mathbf{p}_i)\rangle$, the minimization diagram can be conveniently computed from the intersection of halfspaces bounded by these hyperplanes (Figure 4).

Moreover, it is known [1] that *any* affine Voronoi diagram is a *power diagram* in disguise. We find the explicit equivalence by identifying the respective bisector equations: $\langle \mathbf{x} - \nabla F(\mathbf{p}_i), \mathbf{x} - \nabla F(\mathbf{p}_i)\rangle = \langle \nabla F(\mathbf{p}_i), \nabla F(\mathbf{p}_i)\rangle + 2(F(\mathbf{p}_i) - \langle \mathbf{p}_i, \nabla F(\mathbf{p}_i)\rangle)$, $i = 1, \ldots, n$. It turns out that the first-type Bregman Voronoi diagram is a power diagram of spheres centered at gradient positions $\nabla F(\mathbf{p}_i)$ for weights $r_i^2 = \langle \nabla F(\mathbf{p}_i), \nabla F(\mathbf{p}_i)\rangle + 2(F(\mathbf{p}_i) - \langle \mathbf{p}_i, \nabla F(\mathbf{p}_i)\rangle)$ (possibily imaginary radii). See Figure 5 for an illustration for the exp. loss.

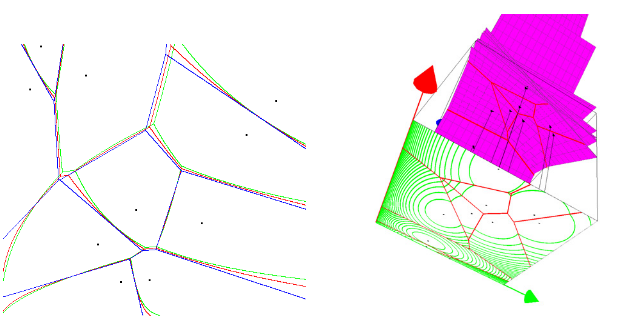

Figure 4: Three types of Bregman Voronoi diagrams for the relative entropy (Kullback-Leibler divergence). First-type affine Bregman Voronoi diagram (blue), second-type Bregman Voronoi diagram (green) and symmetrized Bregman Voronoi diagram (red). First-type Voronoi Bregman diagrams visualized as affine minimization diagrams.

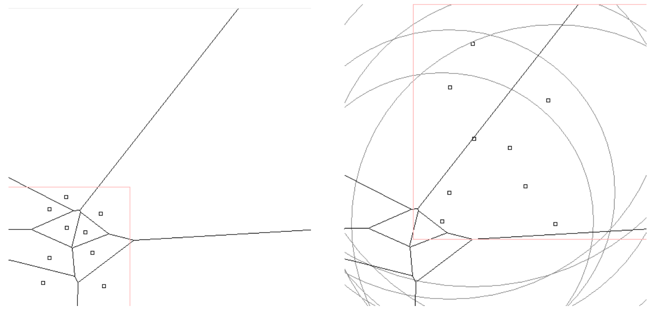

Figure 5: Affine Bregman Voronoi diagrams can be computed as Power diagrams. Illustration for the 2D exponential loss $F(p_x, p_y) = \exp p_x + \exp p_y$: (a) Affine Bregman Voronoi diagram (all cells non-empty), and (b) equivalent Power diagram.

4. STATISTICAL VORONOI DIAGRAMS

A statistical space $\mathcal{X}$ is an abstract space where coordinates of vector points $\boldsymbol{\theta} \in \mathcal{X}$, representing random variables, encode the parameters of statistical distributions. For example, the space $\mathcal{X} = \{[\mu\ \sigma]^T \mid (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+_*\}$ of univariate Normal distributions $\mathcal{N}(\mu, \sigma)$ of mean $\mu \in \mathbb{R}$ and standard deviation $\sigma \in \mathbb{R}^+_*$ is a 2D parametric statistical space, extensively studied in information geometry. A prominent class of distribution families called the *exponential families* $\mathcal{F}_F$ [2] admits the same *canonical* probabilistic distribution function $p(x|\boldsymbol{\theta}) \stackrel{\text{def}}{=} \exp\{\langle \boldsymbol{\theta}, \mathbf{f}(x)\rangle - F(\boldsymbol{\theta}) + C(x)\}$, where $\mathbf{f}(x)$ denotes the sufficient statistics and $\boldsymbol{\theta} \in \mathcal{X}$ represents the *natural parameters*. Space $\mathcal{X}$ is thus called the natural parameter space. $F(\boldsymbol{\theta}) = \log \int_x \exp\{\langle \boldsymbol{\theta}, \mathbf{f}(\mathbf{x})\rangle + C(x)\} \mathrm{d}x$ is referred to as the cumulant function or the log-partition function, and $C(x)$ is a density normalization term. It turns out that the Kullback-Leibler divergence (generator $F(x) = x \log x$, negative Shannon entropy) of any two distributions of the *same* exponential family with respective natural parameters $\boldsymbol{\theta}_p$ and $\boldsymbol{\theta}_q$ is obtained from the Bregman divergence induced by the cumulant function of that family by *swapping* arguments [3]: $\mathrm{KL}(\boldsymbol{\theta}_p||\boldsymbol{\theta}_q) = D_F(\boldsymbol{\theta}_q||\boldsymbol{\theta}_p)$. In this video, we present the Voronoi diagram of univariate Normal distributions as a special case of Bregman Voronoi diagrams for natural parameters ($x = \frac{\mu}{\sigma^2}, y = -\frac{1}{2\sigma^2}$) in the lower half-plane, and cumulant function $F(x,y) = -\frac{x^2}{4y} + \frac{1}{2}\log\frac{-\pi}{y}$.

5. REFERENCES

[1] F. Aurenhammer and R. Klein. Voronoi Diagrams. In *Handbook of Computational Geometry*, pp. 201–290. Elsevier, 2000.

[2] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh. Clustering with Bregman divergences. In *SIAM Data Mining*, pp. 234–245, 2004.

[3] F. Nielsen, J.-D. Boissonnat, and R. Nock. On Bregman Voronoi diagrams. In *ACM-SIAM Symposium on Discrete Algorithms*, pp. 746-755, 2007.