CIMPA Research School : Data Science for Engineering and Technology
Tunis 2019

Bregman divergences
a basic tool for pseudo-metrics building for data structured by physics

6b- Course's take away

Stéphane ANDRIEUX

ONERA - France

Member of the National Academy of Technologies of France

The intention

Try to generalize and extend the Bregman divergence to the processing of data with a rich physical content and strongly mathematically constrained (PDE)

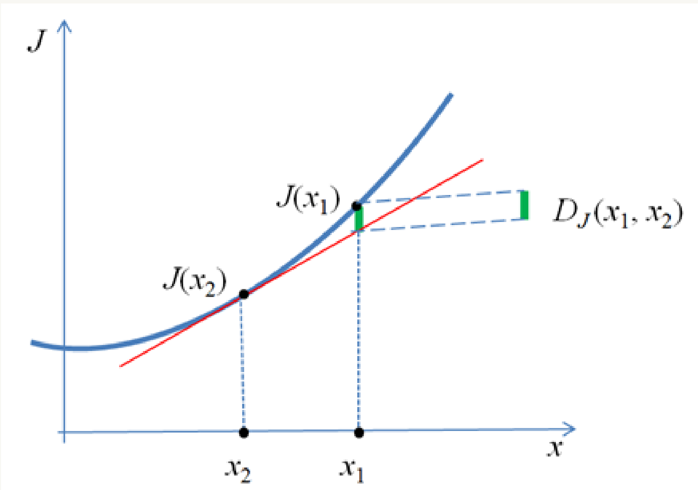
Then introducing and revisiting some techniques and methods used in learning methods :

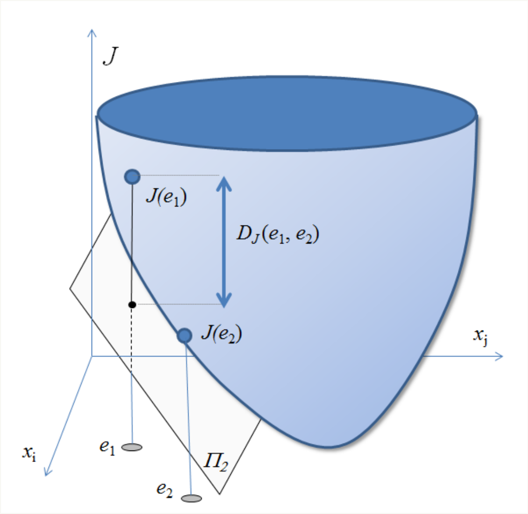
Elementary Computational Geometry – Clustering – POD

Try to overcome the limits brought by the « no triangle inequality » by taking advantage of convexity property and related strong results

And how thermodynamics can help generating « generating function »

The Bregman divergence





$$\boxed{D_J(e_1,e_2) = J(e_1) - J(e_2) - \langle \nabla J(e_2), e_1 - e_2 \rangle}$$

For any J strictly convex $\quad D_J(e_1,e_2) \geq 0 \quad , \quad D_J(e_1,e_2) = 0 \Rightarrow e_1 = e_2$

Not symmetry - No triangle inequality

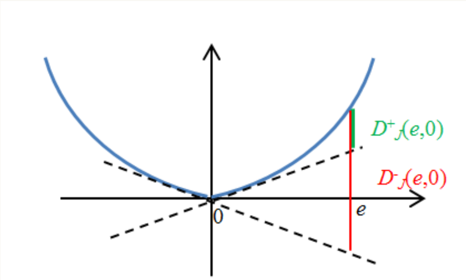
Some are norms (squared Euclidian distance, Mahalanobis distance)
Some existing divergences used in learning are actually BD

Extensions of Bregman divergences

Dealing with non differentiability : extended Bregman divergence

$$D^s{}_J(e_1,e_2) = \min_{p \in \partial J(e_2)} J(e_1) - J(e_2) - \langle p, e_1 - e_2 \rangle \equiv J(e_1) - J(e_2) - \langle \bar{p}_2, e_1 - e_2 \rangle$$

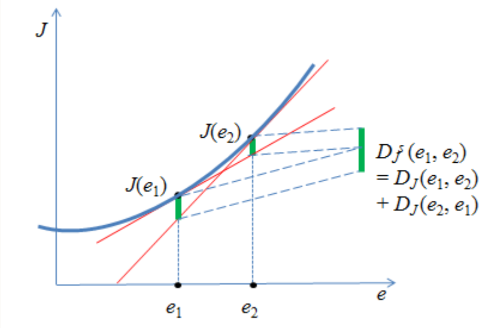
with $$\bar{p}_2 = \underset{p_2 \in \partial J(e_2)}{\arg\min}\ J(e_1) - J(e_2) - \langle p_2, e_1 - e_2 \rangle = \underset{p_2 \in \partial J(e_2)}{\arg\max}\ \langle p_2, e_1 - e_2 \rangle$$



$J(x) = \alpha x^2 + |x|$

Dealing with non symmetry : Symmetrized Bregman Divergence

$$D^s_J(e_1,e_2) = D_J(e_1,e_2) + D_J(e_2,e_1) = \langle \nabla J(e_1) - \nabla J(e_2), e_1 - e_2 \rangle$$



Dealing with pair of (physically) dual variables : Bregman Gaps

$$BG^s_J\left([e_1,p_1],[e_2,p_2]\right) = BG_J\left(e_1,[e_2,p_2]\right) + BG_J\left(e_2,[e_1,p_1]\right) \qquad p_1 \in \partial J(e_1),\ p_2 \in \partial J(e_2)$$

Computational geometry with Bregman divergences

Buy using :

The partial convexity of the Bregman divergence $DJ(.,e)$ for any e

The convexity of the symmetrized Bregman Gap

We can define

The notion of Bregman projection on closed convex sets

The notion of (convex) Bregman ball and Bregman sphere

The notion of equidivergence separation bi-sector (hyperplane)

The notion of pseudo-norm and pseudo-scalar product

Adapting POD and Clustering – Finding Generating functions

By revisiting the methods and carefully separate the various roles of scalar products and norms in standard Euclidian based approaches

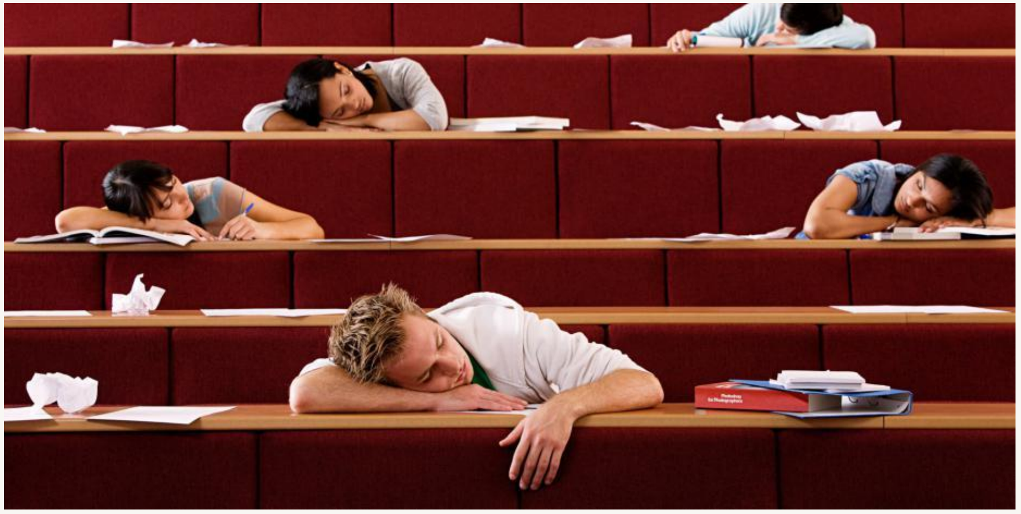
We can built Bregman based POD and Clustering

By taking advantage of

The convex framework of linear symmetric PDE
The existence of various potentials in fluid mechanics
The convex framework of Standard Generalized Materials for the formulation of constitutive equation of materials in solid mechanics

We can built convex generating functions for the Bregman notions

Thanks for your constant attention
and questions



Have a fruitful second week with CIMPA !