Misc: Bregman Divergences

Tuesday, December 10, 2024 3:44 PM

Bregman Divergences

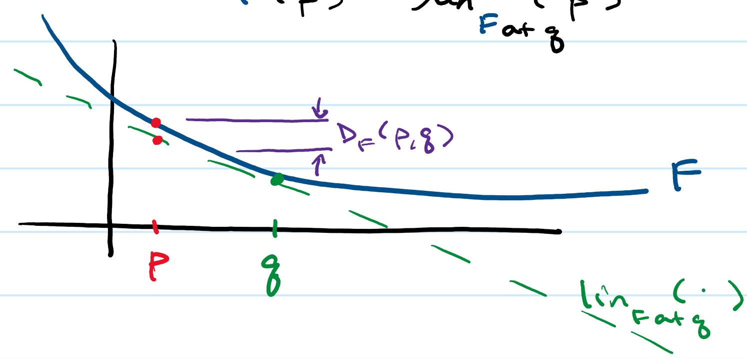
Let F be a strictly convex function (w/ convex domain),

then F induces a Bregman Divergence

or "subgradient" if not differentiable

$$D_F(p,q) := F(p) - F(q) - \langle \nabla F(q),\, p-q \rangle$$

$$= F(p) - \mathrm{lin}_{F\,\mathrm{at}\,q}(p)$$



* A very general class of (almost) metrics: (This is a main reason we care about these)

it's not necessarily symmetric and never satisfies triangle inequality

Properties

① non-negative (by convexity)

② $D_F(p,q) = 0$ iff $p = q$

... see wikipedia for more ...

Examples

① $F(p) = \|p\|^2$ so $\nabla F(p) = 2 \cdot p$

$p \in \mathbb{R}^n$

$$\text{thus } D_F(p,q) = \|p\|^2 - \|q\|^2 - \overbrace{2\langle q, p-q\rangle}^{2\langle q,p\rangle - 2\|q\|^2}$$

$$= \|p\|^2 + \|q\|^2 - 2\langle q,p\rangle$$

$$= \underline{\|p-q\|^2}$$

② $F(p) = \sum_{i=1}^{n} p_i \cdot \log(p_i)$ Negative entropy

$p \in \mathbb{R}^n_+$

$\sum p_i = 1$ (probabilities)

$$\text{then } D_F(p,q) = \underbrace{\sum_i p_i \log\left(\frac{p_i}{q_i}\right)}_{\text{Kullback-Leibler divergence}} - \cancel{\sum_i p_i}^{=1} + \cancel{\sum_i q_i}^{=1} \quad \text{if probability}$$

(hence KL divergence is non-negative)