Cross Entropy under Linear Constraints

Nicholas Barnfield Michael P. Friedlander Tim Hoheisel

2024-04-03

Problem statement

Define the cross-entropy and log-exponential functions by

$$\text{cent}\ (p \mid q) = \sum_{i=1}^{n} p_i \log(p_i / q_i),$$

$$\text{logexp}\ (p \mid q) = \log \sum_{i=1}^{n} q_i \exp(p_i),$$

which form a conjugate pair. We consider the pair of dual optimization problems

$$\min_{p,\, y} \left\{ \text{cent}\ (p \mid q) + \lambda/2 \parallel y \parallel^2 \mid Ax + \lambda y = b \right\}$$

$$\min_{y} \left\{ \text{logexp}\ (A^T y \mid q) + \lambda/2 \parallel y \parallel^2 - \langle b, y \rangle \right\},$$

where A is an m-by-n matrix, b is an m-vector, and λ is a nonnegative regularization parameter.

Implicit approach

Optimality conditions

A pair (p, y) is primal-dual optimal if and only if it satisfies the optimality conditions

$$b = Ax + \lambda y,$$

$$x_j = \frac{1}{\sigma(x)} q_j \exp(\langle a_j, y \rangle), \quad j \in [n],$$

$$\sigma(x) = \sum_{j=1}^{n} q_j \exp(\langle a_j, y \rangle),$$