Comp6211e Homework 3

Assignment date: April 3
Due date: April 17

Theoretical Problems (10 points)

1. (4 points) Consider $x \in C = \{x \in \mathbb{R}^d_+ : \sum_{j=1}^d x_j = 1\}$, and $\mu \in \mathbb{R}^d_+$. Let

$$f(x) = \sum_{j=1}^d x_j \ln \frac{x_j}{\mu_j}.$$

   - Find the conjugate $f^*(y)$.
   - Find the Bregman divergence $D_f(x', x)$ and $D_{f^*}(y, y')$.

2. (3 points) Consider a symmetric positive definite matrix A, and let

$$f(x) = \frac{1}{2} x^\top A x - b^\top x, \quad g(x) = \frac{\lambda}{2}\|x\|_2^2 + \mu\|x\|_1.$$

   Find the Fenchel's dual of $f(x) + g(x)$, and find $\nabla g^*(\alpha)$.

3. (3 points) Consider the matrix regression problem, where the parameter is the matrix $W \in \mathbb{R}^{d\times k}$, and observations $X \in \mathbb{R}^{d\times k}$. Let $\|W\|_*$ be the trace norm of W. Use the SVD decomposition of W to find the solution of

$$\text{prox}(W) = \arg\min_Z \left[\frac{1}{2}\langle W - Z, W - Z\rangle + \lambda\|Z\|_*\right].$$

Programming Problem (10 points)

- Download data in the mnist/ directory (which contains class 1 (positive) versus 7 (negative) from the MNIST data)
- Use the python template "prog_template.py", and implement functions marked with '# implement'.
- Submit your code and outputs. Compare to the theoretical convergence rates in class, and discuss your experimental results.