EE 127 4/15/14
L. El Ghaoui

Homework Assignment #5

Due date: 4/29/14, in class.

Exercise 1 (Some simple optimization problems) Solve the following optimization problems. Make sure to determine an optimal primal solution.

1. Show that, for given scalars α, β,

$$f(\alpha, \beta) \doteq \min_{d>0} \ \alpha d + \frac{\beta^2}{d} = \begin{cases} -\infty & \text{if } \alpha \leq 0 \\ 2|\beta|\sqrt{\alpha} & \text{otherwise.} \end{cases}$$

2. Show that for an arbitrary vector $z \in \mathbb{R}^m$,

$$\|z\|_1 = \min_{d>0} \ \frac{1}{2} \sum_{i=1}^{m} \left(d_i + \frac{z_i^2}{d_i} \right). \tag{1}$$

3. Show that for an arbitrary vector $z \in \mathbb{R}^m$, we have

$$\|z\|_1^2 = \min_{d} \ \sum_{i=1}^{m} \frac{z_i^2}{d_i} \ : \ d > 0, \ \ \sum_{i=1}^{m} d_i = 1.$$

Exercise 2 (Minimizing a sum of logarithms) Consider the following problem:

$$\begin{aligned} p^* = \max_{x \in \mathbb{R}^n} \quad & \textstyle\sum_{i=1}^{n} \alpha_i \ln x_i \\ \text{s.t.} \quad & x \geq 0, \quad \mathbf{1}^\top x = c, \end{aligned}$$

where $c > 0$ and $\alpha_i > 0$, $i = 1, \ldots, n$. Problems of this form arise, for instance, in maximum-likelihood estimation of the transition probabilities of a discrete-time Markov chain. Determine in closed-form a minimizer, and show that the optimal objective value of this problem is

$$p^* = \alpha \ln(c/\alpha) + \sum_{i=1}^{n} \alpha_i \ln \alpha_i,$$

where $\alpha \doteq \sum_{i=1}^{n} \alpha_i$.

Exercise 3 (Logistic Regression) In logistic regression, we are given a set of data points $\{\mathbf{x}_i\}_{i=n}^m$, $\mathbf{x}_i \in \mathbb{R}^n$, and corresponding labels $\{y_i\}_{i=1}^n$ where $y_i \in \{0,1\}$. Our goal is to solve the following optimization problem

$$\min_{\mathbf{w},b} \sum_{i=1}^{m} -y_i(\mathbf{w}^T\mathbf{x}_i + b) + \log\left(1 + \exp(\mathbf{w}^T\mathbf{x}_i + b)\right) \tag{2}$$

1. By slightly modifying the definition of $\mathbf{x}$, explain how we can re-write the problem as

$$\min_{\mathbf{w}} \sum_{i=1}^{m} -y_i(\mathbf{w}^T\mathbf{x}_i) + \log\left(1 + \exp(\mathbf{w}^T\mathbf{x}_i)\right) \tag{3}$$

2. Is (3) a convex optimization problem? Explain.

3. Show that we can write the gradient of the objective function in (3) in the form

$$\nabla f(\mathbf{w}) = \mathbf{M}^T(\mathbf{z} - \mathbf{u})$$

   for a matrix $\mathbf{M}$ and vectors $\mathbf{z}, \mathbf{u}$ which you will determine.

4. Show that we can write the Hessian of the objective function in (3) in the form

$$\nabla^2 f(\mathbf{w}) = \mathbf{M}^T\mathbf{D}\mathbf{M}$$

   for the same matrix $\mathbf{M}$ as in problem 3 and where $\mathbf{D} = \mathbf{Diag}(\mathbf{v})$ for $\mathbf{v}$ which you will determine.

5. Write the gradient method update for solving (3).

6. Implement the gradient method for solving (3) with the data included in the file `adult_data.mat`.

   (a) Try at least three different step sizes (e.g. α small, $\alpha_k = 1/k$, $\alpha_k = 1/\sqrt{k}$, or α of your choice) .

   (b) Provide diagnostics for each step size. For example, provide average iteration time, the number of iterations until convergence (or if it reaches the maximum number of iterations), and time until convergence.