Management Sciences Topics: Convex Optimization
Midterm

1. Problem 1.

   (a)
   $$f(x,y) = \log(x^2 + xy + y^2 + 1),\ x, y \in \mathbb{R}.$$
   Not convex.

   (b)
   $$f(x) = \exp(-\|Ax + b\|_2^2)$$
   Not convex.

   (c)
   $$f(x) = \exp(\|Ax + b\|_2^2)$$
   Convex. Let $f(x) = h \circ g(x)$, where $h(x) = \exp(x)$ is convex and increasing, and $g(x) = \|Ax + b\|_2^2$ is convex. Hence f is convex.

2. Problem 2.

   The Lagrange dual function is
   $$g(\lambda, \nu) = \inf_{x \in \mathbb{R}_+^n} \sum_{i=1}^{n} x_i \ln(x_i/y_i) + \sum_{i=1}^{n} \lambda_i (x_i - b_i) + \sum_{i=1}^{n} \lambda_i' (a_i - x_i) + \nu(\sum_{i=1}^{n} x_i - 1),$$
   where $\lambda = (\lambda_i, \lambda_i') \in \mathbb{R}^{2n}$. Then
   $$\frac{\partial}{\partial x_i} g = \ln(x_i/y_i) + y_i + \lambda_i - \lambda_i' + \nu,$$
   so in order to get the minimum value for g,
   $$x_i = y_i \exp(-y_i - \lambda_i + \lambda_i' - \nu).$$

Then

$$
\begin{aligned}
g(\lambda, \nu) = &\sum_{i=1}^{n} y_i(-y_i - \lambda_i + \lambda_i' - \nu)\exp(-y_i - \lambda_i + \lambda_i' - \nu) \\
&+ \sum_{i=1}^{n} \lambda_i(y_i \exp(-y_i - \lambda_i + \lambda_i' - \nu) - b_i) \\
&+ \sum_{i=1}^{n} \lambda_i'(a_i - y_i \exp(-y_i - \lambda_i + \lambda_i' - \nu)) \\
&+ \nu(\sum_{i=1}^{n} y_i \exp(-y_i - \lambda_i + \lambda_i' - \nu) - 1) \\
= &- \sum_{i=1}^{n} y_i^2 e^{-y_i - \lambda_i + \lambda_i' - \nu} - \sum_{i=1}^{n} \lambda_i b_i + \sum_{i=1}^{n} \lambda_i' a_i - \nu.
\end{aligned}
$$

Hence, the dual problem is

$$
\begin{aligned}
\text{maximize} \quad & -\sum_{i=1}^{n} y_i^2 e^{-y_i - \lambda_i + \lambda_i' - \nu} - \sum_{i=1}^{n} \lambda_i b_i + \sum_{i=1}^{n} \lambda_i' a_i - \nu \\
\text{subject to} \quad & \lambda_i \geq 0, \lambda_i' \geq 0.
\end{aligned}
$$

The KKT condition is

$$
\begin{aligned}
x_i - b_i \leq 0, &\quad i = 1, \ldots, n \\
-x_i + a_i \leq 0, &\quad i = 1, \ldots, n \\
\lambda_i \geq 0, &\quad i = 1, \ldots, n \\
\lambda_i' \geq 0, &\quad i = 1, \ldots, n \\
\lambda_i(x_i - b_i) = 0, &\quad i = 1, \ldots, n \\
\lambda_i'(-x_i + a_i) = 0, &\quad i = 1, \ldots, n \\
\ln(x_i/y_i) + y_i + \lambda_i - \lambda_i' + \nu_i = 0, &\quad i = 1, \ldots, n
\end{aligned}
$$

(I omitted the $*$ for each variable $x_i^*, \lambda_i^*, \lambda_i'^*, \nu_i^*$.)

3. Problem 3.

Proof. By definition,

$$
\mathrm{Prox}_h(z) = \arg\min_{x \in \mathbb{R}^d} \frac{1}{2}\|x - z\|_2^2 + \lambda\|x\|_2.
$$

Let

$$
f(x) = \frac{1}{2}\|x - z\|_2^2 + \lambda\|x\|_2,
$$

then by definition of subgradients,

$$\partial f = \begin{cases} x - z + \lambda \dfrac{x}{\|x\|_2}, & x \neq 0, \\ -z + \lambda \bar{B}(0,1), & x = 0, \end{cases}$$

where $\bar{B}(0,1)$ is the closed unit ball. Then

(a) If $\|z\|_2 > \lambda$, then $0 \notin -z + \lambda \bar{B}(0,1)$. In this case, the optimal x satisfies $x \neq 0$ and

$$x - z + \lambda \frac{x}{\|x\|_2} = 0.$$

By solving this, we get

$$x^* = (\|z\|_2 - \lambda)\frac{z}{\|z\|_2}.$$

Since $x^* \neq 0$, this holds when $\|z\|_2 > \lambda$.

(b) When $\|z\| \leq \lambda$, $0 \in -z + \lambda \bar{B}(0,1)$, while there's no such x s.t. $x - z + \lambda \frac{x}{\|x\|_2} = 0$. Hence $x^* = 0$. Thus we have proved the statement.

□

4. Problem 4.

The code for this problem are `accelerated_proximal_gradient.m` and `elastic_net_regularized_logistic_regression.m`. The problem is formatted as

$$f(x) = g(x) + h(x),$$

where

$$g(x) = \frac{1}{n}\sum_{i=1}^{n} \log(1 + \exp(-b_i a_i^\top x) + \frac{\lambda_2}{2}\|x\|_2^2,$$

and $h(x) = \lambda_1 \|x\|_1$, so as to guarantee the strongly convexity of g.

However, if I choose $x_0 = 0$, I have $x_1 = x_0$ exactly, so the iteration cannot continue, which does not agree with the properties.

5. Problem 5.

The code for this problem are `stochastic_subgradient.m` and `sparse_hinge_loss.m`.

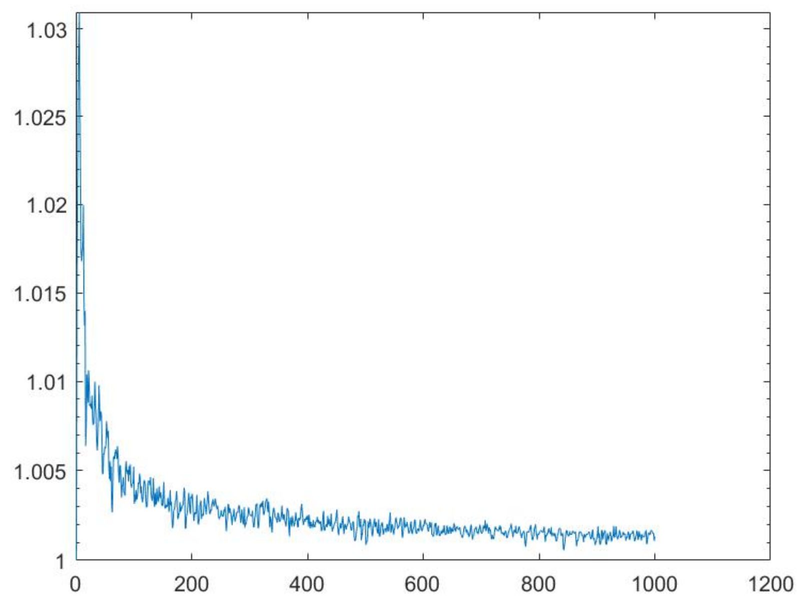

Figure 1: Convergence plot for SSG