Math 582g Convex Optimization : Final report

Karthik Mohan

March 20, 2008

1 Problem 4.61 - Optimization with logistic model

A random variable X $\in \{0, 1\}$ satisfies

$$prob(X = 1) = p = \frac{e^{a^T x+b}}{1 + e^{a^T x+b}}$$

where x $\in R^n$ is a vector of variables that affect the probability, and a and b are known parameters. We can think of X = 1 as the event that a consumer buys a product, and x as a vector of variables that affect the probability, e.g. advertising effort, retail price, discounted price, packaging expense, and other factors. The variable x, which we are to optimize over, is subject to a set of linear constraints, $Fx \leq g$. Formulate the following problems as convex optimization problems.

1.1 Problem 4.61, Part a

(a) *Maximizing buying probability.* The goal is to choose x to maximize p. The problem is to maximize buying probability,

$$p = \frac{e^{a^T\, x+b}}{1 + e^{a^T\, x+b}}$$

subject to $Fx \leq g$. Let $f(y) = \frac{e^y}{1+e^y}$
Then, $f(y)$ is increasing in y $\forall$ y. (This is because $f^{'}(y) = \frac{e^y}{(1+e^y)^2} \geq 0$)
$\therefore$ maximizing $f(y)$ subject to given constraints is $\equiv$ to maximizing
y subject to the given constraints. Since $p = f(a^T\, x + b)$, the formulation for this problem is an LP and given by:

$$\text{Maximize } a^T\, x + b \text{ subject to } \quad Fx \leq g$$

The dual of the formulation is given by:

$$\text{Minimize } y^T g \text{ subject to } \quad F^T y = a,$$
$$y >= 0$$

1.2 Problem 4.61, Part b

Maximizing expected profit. Let $c^T x + d$ be the profit derived from selling the product,which we assume is positive for all feasible x. The goal is to maximize the expected profit, which is $p(c^T x + d)$.
The problem is to maximize the expected profit.
The expected profit is given by,

E(profit) = P(customer buys product)× (Profit when customer buys product)

+

P(customer doesn't buy product)×(Profit when customer doesn't buy product)

$$\therefore \text{ E(profit)} = \text{p} \times \; (c^T x + d) \; = g(x)$$

Check for convexity: Let $h(s,t) = \frac{s}{1+e^{-t}}$.
Note that, $g(x) = h(c^T\, x + d,\, a^T\, x + b)$
To check if $g(x)$ is convex, we compute the gradient and hessian of $h(s,t)$

The gradient is given by,

$$\bigtriangledown h(s,t) = \begin{bmatrix} \frac{1}{(1+e^{-t})} \\ \frac{se^{-t}}{(1+e^{-t})} \end{bmatrix}$$

and the hessian is given by,

$$\bigtriangledown^2 h(s,t) = \begin{bmatrix} 0 & \frac{e^{-t}}{(1+e^{-t})^2} \\ \frac{e^{-t}}{(1+e^{-t})^2} & \frac{se^{-t}(1-e^{-t})}{(1+e^{-t})^3} \end{bmatrix}$$

Clearly, the hessian is not positive semidefinite. The hessian is also not negative semidefinite since the second principal minor, which equals the determinant of the hessian, is less than zero. Hence, $h(s,t)$ is neither convex nor concave.

Check for quasi-convexity: $g(x) = h(c^T x + d, a^T x + b)$ can be tested for quasi-convexity. If $g(x)$ is quasi-convex or quasi-concave, we can use quasi-convex optimization to determine the local/global optima through techniques mentioned in section 4.2.5 of the textbook (Convex Optimization by Stephen Boyd).
Now, for any $y \in R$, $f(y)$ is increasing and hence is quasi-convex.
$-f(y)$ is decreasing and hence quasi-convex.

$\therefore$ $f(y)$ is both quasi-convex and quasi-concave. Consider two feasible points x_1, x_2. Then $c^T x_1 + d \geq 0, c^T x_2 + d \geq 0$ (given).
Also, let $y_1 = a^T x_1 + b, y_2 = a^T x_2 + b$, then

$$f(y_1) \leq \alpha \text{ implies}$$

$$\frac{1}{1 + e^{-(a^T x_1 + b)}} \leq \alpha \tag{1}$$

and

$$f(y_2) \leq \alpha \text{ implies}$$

$$\frac{1}{1 + e^{-(a^T x_2 + b)}} \leq \alpha \tag{2}$$

and

$$f(\lambda y_1 + (1 - \lambda) y_2) \leq \alpha \text{ implies}$$

$$\frac{1}{1 + e^{-(\lambda(a^T x_1 + b) + (1 - \lambda)(a^T x_2 + b))}} \leq \alpha \tag{3}$$

(1) $\times$ $(c^T x_1 + d)$, (2) $\times$ $(c^T x_2 + d)$ implies

$$g(x_1) = \frac{c^T x_1 + d}{1 + e^{-(a^T x_1 + b)}} \leq \alpha(c^T x_1 + d) \leq \alpha' \tag{4}$$

and

$$g(x_2) = \frac{c^T x_2 + d}{1 + e^{-(a^T x_2 + b)}} \leq \alpha(c^T x_2 + d) \leq \alpha' \tag{5}$$

where, $\alpha' = \alpha max(c^T x_1 + d, c^T x_2 + d)$

(3) $\times$ $(\lambda(c^T x_1 + d) + (1 - \lambda)(c^T x_2 + d))$ implies

$$\frac{\lambda(c^T x_1 + d) + (1-\lambda)(c^T x_2 + d)}{1 + e^{-(\lambda(a^T x_1 + b) + (1-\lambda)(a^T x_2 + b))}}$$
$$\le \alpha(\lambda(c^T x_1 + d) + (1-\lambda)(c^T x_2 + d))$$
$$\le \lambda(\alpha(c^T x_1 + d)) + (1-\lambda)(\alpha(c^T x_2 + d))$$
$$\le \lambda(\alpha^{'}) + (1-\lambda)(\alpha^{'})$$
$$\le \alpha^{'}$$

i.e.,

$$g(\lambda x_1 + (1-\lambda)x_2) = \frac{\lambda(c^T x_1 + d) + (1-\lambda)(c^T x_2 + d)}{1 + e^{-(\lambda(a^T x_1 + b) + (1-\lambda)(a^T x_2 + b))}} \le \alpha^{'} \tag{6}$$

(4),(5) and (6) implies $g(x) = h(c^T x + d,\, a^T x + b)$ is quasi-convex.

(This is because, for every $\alpha^{'}, x_1, x_2,$ there is a corresponding α given by,

$$\alpha = \frac{\alpha^{'}}{max(c^T x_1 + d, c^T x_2 + d)})$$

Similarly it can be shown that $-g(x)$ is quasi-convex, which implies that $g(x)$ is quasi-linear.

Hence the quasi-convex formulation for this problem is:

$$\text{Minimize} - g(x) \text{ subject to}$$
$$Fx \le g$$

The equivalent formulation in terms of the function $h(s,t)$ is

$$\text{Minimize} - h(s,t) \text{ subject to}$$
$$Fx \le g,$$
$$s - (c^T x + d) = 0,$$
$$t - (a^T x + b) = 0$$

The dual of this formulation is:

$$D : sup_{y,u,v} inf_{x,s,t} \left(-h(s,t) + y^T(Fx - g) + u(s - (c^T x + d)) + v(t - (a^T x + b)) + \delta_{R_m^+}(y)\right)$$

Rearranging terms, we get,

$$D : sup_{y,u,v,y\ge 0} \left(-y^T g - ud - vb + inf_x \left(y^T Fx - uc^T x - va^T x + inf_{s,t}\left(\frac{-s}{1 + e^{-t}} + us + vt\right)\right)\right)$$

implies

$$D : sup_{y,u,v,y\geq 0,u\geq 0,u\leq 1}\left(-y^T g - ud - vb + v\log\left(\frac{u}{1-u}\right) + inf_x\left(y^T Fx - uc^T x - va^T x\right)\right)$$

$$\text{implies } D : sup_{y,u,v,y\geq 0,u\geq 0,u\leq 1}\left(-y^T g - ud - vb + v\log\left(\frac{u}{1-u}\right)\right)$$

subject to

$$y^T F = uc^T + va^T$$