ELEC 5470 HW 4

TANG Jiawei
20672880

Problem: Lasso

$$\text{minimize } \|y - X\beta\|_2^2 + \lambda \|\beta\|_1$$

The problem can be reformulated as

$$\text{minimize } (y - X\beta)^T (y - X\beta) + \lambda 1^T t$$

$$\text{s.t.} \quad \beta - t \le 0$$

$$-\beta - t \le 0$$

With log Barrier. The problem can be approximated as

$$\text{minimize } (y - X\beta)^T (y - X\beta) + \lambda 1^T t + \phi(\beta)$$

$$\text{with } \phi(\beta) = -\frac{1}{\alpha}\left[\sum_{i=1}^{p} \log(t_i - \beta_i) + \sum_{i=1}^{p} \log(\beta_i + t_i)\right]$$

$$\alpha \ge 0$$

$$\therefore L(\beta, t) = (y - X\beta)^T (y - X\beta) + \lambda 1^T t - \frac{1}{\alpha}\left(\sum_{i=1}^{p} \log(t_i - \beta_i) + \sum_{i=1}^{p} \log(\beta_i + t_i)\right)$$

$$\nabla_\beta L(\beta, t) = 2X^T X\beta - 2X^T y + \frac{1}{\alpha}\left[\frac{2\beta_1}{t_1^2 - \beta_1^2}, \frac{2\beta_2}{t_2^2 - \beta_2^2}, \cdots, \frac{2\beta_p}{t_p^2 - \beta_p^2}\right]^T$$

$$\nabla_t L(\beta, t) = \lambda 1 - \frac{1}{\alpha}\left[\frac{2t_1}{t_1^2 - \beta_1^2}, \frac{2t_2}{t_2^2 - \beta_2^2}, \cdots, \frac{2t_p}{t_p^2 - \beta_p^2}\right]^T$$

$\therefore$ The gradient is:

$$\nabla L = [\nabla_\beta L(\beta, t), \nabla_t L(\beta, t)]^T$$

For the hessian matrix of $L(\beta, t)$

$$H_L = \begin{bmatrix} \frac{d^2 L}{d\beta^2} & \frac{d^2 L}{d\beta\, dt} \\ \frac{d^2 L}{dt\, d\beta} & \frac{d^2 L}{dt^2} \end{bmatrix}$$

with

$$\frac{d^2 L}{d^2\beta^2} = X^T X + \frac{1}{\alpha} \operatorname{diag}\left[\frac{2(t_1^2+\beta_1^2)}{(t_1^2-\beta_1^2)}, \ \cdots, \ \frac{2(t_p^2+\beta_p^2)}{(t_p^2-\beta_p^2)} \right]$$

$$\frac{d^2 L}{d\beta\, dt} = -\frac{1}{\alpha} \operatorname{diag}\left[\frac{4\beta_1 t_1}{(t_1^2-\beta_1^2)^2}, \ \cdots, \ \frac{4\beta_p t_p}{(t_p^2-\beta_p^2)^2} \right]$$

$$\frac{d^2 L}{dt\, d\beta} = -\frac{1}{\alpha} \operatorname{diag}\left[\frac{4\beta_1 t_1}{(t_1^2-\beta_1^2)^2}, \ \cdots, \ \frac{4\beta_p t_p}{(t_p^2-\beta_p^2)^2} \right]$$

$$\frac{d^2 L}{dt^2} = \frac{1}{\alpha} \operatorname{diag}\left[\frac{2(t_1^2+\beta_1^2)}{(t_1^2-\beta_1^2)^2}, \ \cdots, \ \frac{2(t_p^2+\beta_p^2)}{(t_p^2-\beta_p^2)^2} \right]$$

The figures and optimal solution are showed in last 2 pages

Problem 2.

1) The problem is not a convex problem as the equality constraint is not affine function.

Construction $u(x, x^k) = (x-x^k)^H M^+ (x-x^k) + x^H M x$

with ① $u(x^k, x^k) = x^{kH} M x^k = f(x^k) \quad \forall x^k \in \mathrm{dom} f$

② $u'(x^k, x^k) = f'(x^k) \quad \forall x^k \in \mathrm{dom} f$

we should set a proper M^+ to make sure

③ $u(x, x^k) \geq f(x) \quad \forall x, x^k \in \mathrm{dom} f$

$\therefore M^+ \succeq 0$, which indicates M^+ should be positive semidefinite.

$\because u(x, x^k) = (x-x^k)^H M^+ (x-x^k) + x^H M x$

$= x^H M^+ x + x^{kH} M^+ x^k - x^H M^+ x^k - x^{kH} M^+ x + x^H M x$

$= x^H (M^+ + M) x + x^{kH} M^+ x^k - x^H M^+ x^k - x^{kH} M^+ x$

here $x^{kH} M^+ x^k$ is constant

Let $M^+ + M = \lambda_{max} I$ where λ_{max} is the largest eigenvalue of M. By SVD $M^+ = U^H (\lambda_{max} I - \Sigma) U \succeq 0$

Then $x^H (M^+ + M) x = \lambda_{max} x^H x$

with $\|x_i\| = 1$, $x^H (M^+ + M) x = n \lambda_{max}$

which is constant

$\therefore$ minimize $U(x, x_k)$

s.t. $\|x_i\| = 1$

is changed to

minimize $-x^{kH} M^+ x - x^H M^+ x^k$

s.t. $|x_i| = 1$

which is

minimize $-2\mathrm{Re}(x^H M^+ x^k)$

s.t. $|x_i| = 1$

Let $y = -M^+ x^k$

The above problem should has closed form solution

with $x_i^k = e^{+j \angle y_i}$ $\forall i = 1, 2, \cdots, n$

The MM alogrithm should be

init $x^0 \in \mathrm{dom} f$, $k=0$

repeat

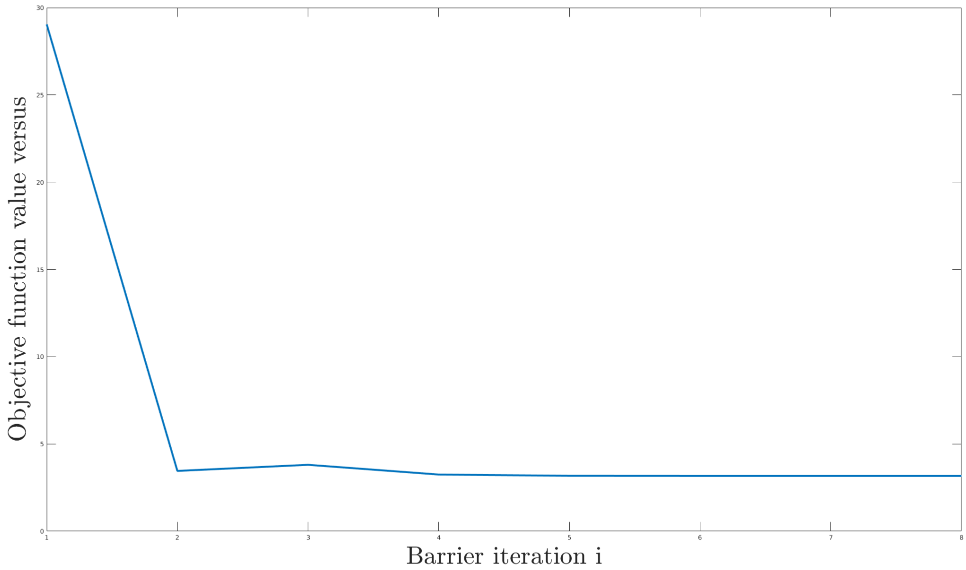
- $y \leftarrow -(\lambda_{max} I - M) x^k$
- $x_i^{k+1} \leftarrow e^{+j \angle y_i}$
- $k \leftarrow k+1$

until converge

return x^k

The figure and optimal solution are showed in last 2 pages.

Problem figure 1

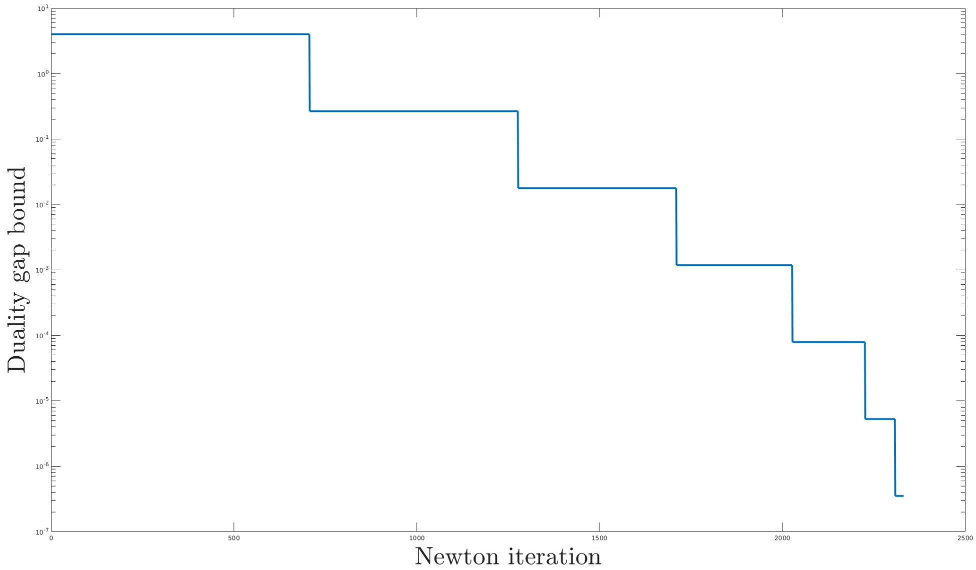


with optimal value $f(\beta^*) = 3.165004206118524$

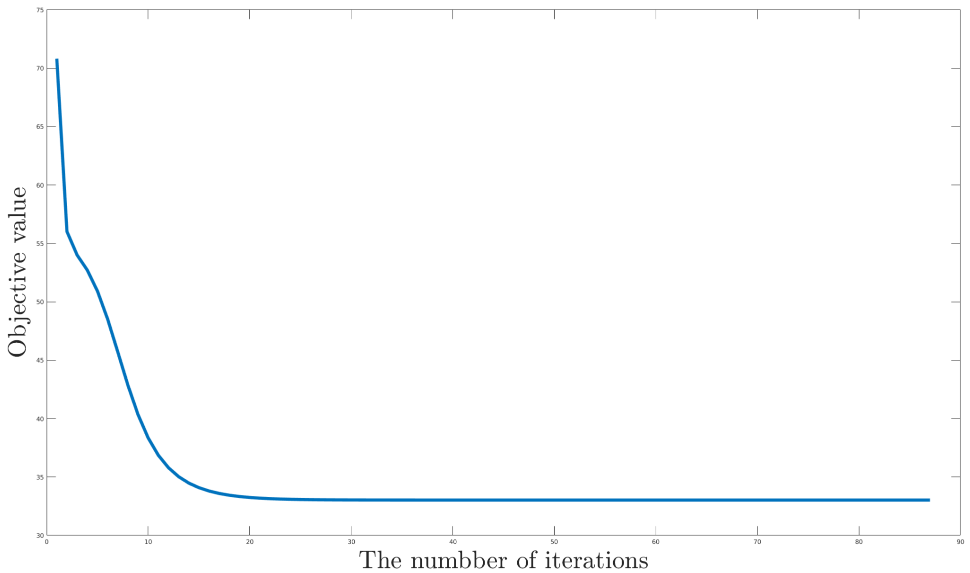
and optimal solution

$\beta^* = [0.000392319376974944$
0.00283173170257793
0.996181204516745
0.000390092311648555
7.01355415531076
0.00313297094830546
0.0119443590635569
0.00104687021823952
0.00213998063005172
$3.00179019836659]$

problem 1 figure 2



problem 2 figure 1



with optimal value = 33.025729646836055