

Lun Zhang

National Key Laboratory for Novel Software Technology Nanjing University Nanjing

ZHANGLJ LAMDA\_NJU\_EDU\_CN

China

Jinfeng Yin

IBM Thomas J. Watson Research Center Yorktown Heights NY 10598 USA

JINFENGY US\_IBM\_COM

Rongjin Jin

Department of Computer Science and Engineering Michigan State University East Lansing MI 48824 USA

RONGJIN CSE\_MSU\_EDU

## A Proof of Lemma 6

We consider the following general optimization problem

$$\min_{\|\mathbf{x}\|_2 \leq 1} -\mathbf{x}^\top \mathbf{y} + \gamma \|\mathbf{x}\|_1.$$

Before we proceed we need the following lemma.

**Lemma 6.** *The solution to the optimization problem*

$$\min_x \frac{1}{2}(x - y)^2 + \gamma|x|$$

is given by

$$P_\gamma(y) = \begin{cases} 0, & \text{if } |y| \leq \gamma; \\ \text{sign}(y)(|y| - \gamma), & \text{otherwise.} \end{cases}$$

where  $P_\gamma(\cdot)$  is the soft-thresholding operator defined in (7) (Donoho, 1995).

The proof of Lemma 6 can be found in Duchi Singer

9. Based on the above lemma it is easy to verify that

$$\min_x \frac{1}{2}(x - y)^2 + \gamma|x| = \begin{cases} \frac{y^2}{2}, & \text{if } |y| \leq \gamma; \\ \gamma|y| - \frac{\gamma^2}{2}, & \text{otherwise.} \end{cases} \quad (6)$$

First we consider the case  $\|y\|_\infty \leq \gamma$ . Then it is easy to verify that

$$\mathbf{0} \in \underset{\mathbf{x}}{\text{argmin}} -\mathbf{x}^\top \mathbf{y} + \gamma \|\mathbf{x}\|_1.$$

Since  $\|\mathbf{0}\|_2 \leq 1$   $\mathbf{0}$  is also an optimal solution to

Next we consider the case  $\|y\|_\infty > \gamma$ . Following the standard analysis of convex optimization Boyd Vandenberghe 4 the Lagrange dual

function  $g(\mu)$  of is given by

$$\begin{aligned} g(\mu) &= \min_{\mathbf{x}} -\mathbf{x}^\top \mathbf{y} + \gamma \|\mathbf{x}\|_1 + \mu(\|\mathbf{x}\|_2^2 - 1) \\ &= \min_{\mathbf{x}} 2\mu \left( \frac{1}{2} \left\| \mathbf{x} - \frac{\mathbf{y}}{2\mu} \right\|_2^2 + \frac{\gamma}{2\mu} \|\mathbf{x}\|_1 \right) - \frac{\|\mathbf{y}\|_2^2}{4\mu} - \mu \\ &= 2\mu \left( \sum_i \min_{x_i} \frac{1}{2} \left( x_i - \frac{y_i}{2\mu} \right)^2 + \frac{\gamma}{2\mu} |x_i| \right) - \frac{\|\mathbf{y}\|_2^2}{4\mu} - \mu \\ &\stackrel{6}{=} 2\mu \left( \sum_{i:|y_i| \leq \gamma} \frac{y_i^2}{8\mu^2} + \sum_{i:|y_i| > \gamma} \left( \frac{\gamma|y_i|}{4\mu^2} - \frac{\gamma^2}{8\mu^2} \right) \right) \\ &\quad - \frac{\|\mathbf{y}\|_2^2}{4\mu} - \mu \\ &= \sum_{i:|y_i| > \gamma} \left( \frac{\gamma|y_i|}{2\mu} - \frac{\gamma^2}{4\mu} - \frac{y_i^2}{4\mu} \right) - \mu \\ &= - \frac{\sum_{i:|y_i| > \gamma} (|y_i| - \gamma)^2}{4\mu} - \mu = - \frac{\|P_\gamma(\mathbf{y})\|_2^2}{4\mu} - \mu. \end{aligned}$$

So the Lagrange dual problem is

$$\max_{\mu \geq 0} - \frac{\|P_\gamma(\mathbf{y})\|_2^2}{4\mu} - \mu$$

and the optimal dual solution is

$$\mu_* = \frac{\|P_\gamma(\mathbf{y})\|_2}{2}.$$

Following the standard analysis Boyd Vandenberghe

4 Section -- the optimal primal solution is

$$\mathbf{x}_* = \underset{\mathbf{x}}{\text{argmin}} \frac{1}{2} \left\| \mathbf{x} - \frac{\mathbf{y}}{2\mu_*} \right\|_2^2 + \frac{\gamma}{2\mu_*} \|\mathbf{x}\|_1$$

$\stackrel{\text{Lemma 6}}{=} \frac{1}{\|P_\gamma(\mathbf{y})\|_2} P_\gamma(\mathbf{y}).$

**B. Proof of Lemma 2**

We first consider the case  $\text{sign}(\mathbf{x}_k^\top \mathbf{u}) = 1$  i.e.

$$\mathbf{x}_k^\top \frac{\mathbf{u}}{\|\mathbf{u}\|_2} > \delta_k.$$

Then we have

$$\begin{aligned} \mathbf{x}_*^\top \frac{\mathbf{u}}{\|\mathbf{u}\|_2} &= \mathbf{x}_k^\top \frac{\mathbf{u}}{\|\mathbf{u}\|_2} + (\mathbf{x}_* - \mathbf{x}_k)^\top \frac{\mathbf{u}}{\|\mathbf{u}\|_2} \\ &> \delta_k - \|\mathbf{x}_* - \mathbf{x}_k\|_2 \stackrel{\Delta}{\geq} 0. \end{aligned}$$

Thus

$$\text{sign}(\mathbf{x}_*^\top \mathbf{u}) = \text{sign}\left(\mathbf{x}_*^\top \frac{\mathbf{u}}{\|\mathbf{u}\|_2}\right) = 1 = \text{sign}(\mathbf{x}_k^\top \mathbf{u})$$