STRONG RULES FOR EFFICIENT LASSO COMPUTATIONS AND THE BASIL ALGORITHM

NOTES BY PARKER KNIGHT

1. Strong rules for the LASSO

1.1. **Subgradients.** Let $f : \mathbb{R}^m \to \mathbb{R}$ be convex.

Recall the following first order condition: if f is differentiable, then

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \forall x, y \in \mathrm{dom}(f)$$

What if f is not differentiable? This motivates the following definition: call $g \in \mathbb{R}^m$ a *subgradient* of f at x iff

$$f(y) \geq f(x) + g^T (y - x) \quad \forall x, y \in \mathrm{dom}(f)$$

The subdifferential of f at x, denote $\partial f(x)$ is the set of all subgradients. The following facts will be useful:

(1) If f is differentiable, then $\partial f(x) = \{\nabla f(x)\}$
(2) For $\alpha_1, \alpha_2 \geq 0$, then $\partial \left[\partial \alpha_1 f_1(x) + \alpha_2 f_2(x)\right] = \alpha_1 \partial f_1(x) + \alpha_2 \partial f_2(x)$
(3) x^* minimizes f iff $0 \in \partial f(x^*)$

where we define set addition as $A + B = \{a + b | a \in A, b \in B\}$.

1.2. **The LASSO.**

1.3. **Strong rules.**

2. BASIL

References