## Machine Learning – Brett Bernstein

## Week 2 Pre-Lecture: Concept Check Exercises

## Optimization Prerequisites for Lasso

## L1 and L2 Regularization

1. Given  $a \in \mathbb{R}$  we define  $a^+, a^-$  as follows:

$$a^+ = \begin{cases} a & \text{if } a \ge 0, \\ 0 & \text{otherwise,} \end{cases}$$
 and  $a^- = \begin{cases} -a & \text{if } a < 0, \\ 0 & \text{otherwise.} \end{cases}$ 

We call  $a^+$  the positive part of a and  $a^-$  the negative part of a. Note that  $a^+, a^- \ge 0$ .

- (a) Give an expression for a in terms of  $a^+, a^-$ .
- (b) Give an expression for |a| in terms of  $a^+, a^-$ . For  $x \in \mathbb{R}^d$  define  $x^+ = (x_1^+, \dots, x_d^+)$  and  $x^- = (x_1^-, \dots, x_d^-)$ .
- (c) Give an expression for x in terms of  $x^+, x^-$ .
- (d) Give an expression for  $||x||_1$  without using any summations or absolute values. [Hint: Use  $x^+, x^-$  and the vector  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^d$ .]
- 2. Let  $f: \mathbb{R} \to \mathbb{R}$  and  $S \subseteq \mathbb{R}$ . Consider the two optimization problems

minimize
$$_{x \in \mathbb{R}}$$
  $|x|$  minimize $_{a,b \in \mathbb{R}}$   $a+b$  subject to  $f(x) \in S$  and subject to  $f(a-b) \in S$   $a,b > 0$ .

Solve the following questions.

- (a) If x in the first problem satisfies  $f(x) \in S$  show how to quickly compute (a, b) for the second problem with a + b = |x| and  $f(a b) \in S$ .
- (b) If a, b in the second problem satisfy  $f(a b) \in S$ , show how to quickly compute an x for the first problem with  $|x| \le a + b$  and  $f(x) \in S$ .
- (c) Assume x is a minimizer for the first problem with minimum value  $p_1^*$  and (a, b) is a minimizer for the second problem with minimum  $p_2^*$ . Using the previous two parts, conclude that  $p_1^* = p_2^*$ .
- 3. Let  $f: \mathbb{R}^d \to \mathbb{R}$ ,  $S \subseteq \mathbb{R}$  and consider the following optimization problem:

$$\begin{array}{ll}
\text{minimize}_{x \in \mathbb{R}^d} & ||x||_1\\ 
\text{subject to} & f(x) \in S,
\end{array}$$

where  $||x||_1 = \sum_{i=1}^d |x_i|$ . Give a new optimization problem with a linear objective function and the same minimum value. Show how to convert a solution to your new problem into a solution to the given problem. [Hint: Use the previous two problems.]

1