CS 6316 Homework 4

James Harbour
gtr8rh@virginia.edu

March 25, 2024

1 Problem 1

Show how to cast the ERM problem of linear regression with respect to the absolute value loss function, $l(h,(x,y)) = |h(x) - y|$, as a linear program; namely, show how to write the problem

$$\min_{w} \sum_{i=1}^{m} |\langle w, x_i \rangle - y_i|$$

as a linear program. *Hint*: Start with proving that for any $c \in \mathbb{R}$,

$$|c| = \min_{a \geq 0} a \text{ s.t } c \leq a \text{ and } c \geq -a.$$

You can use the hint directly without proving it.

Proof. First we rewrite the minimization problem in terms of an inner product and some constraints

$$\min_{w \in \mathbb{R}^d} \sum_{i=1}^{m} |\langle w, x_i \rangle - y_i| = \begin{cases} \min_{u=(w_1,\ldots,w_d,a_1,\ldots,a_d) \in \mathbb{R}^{d+m}} \sum_{j=1}^{d} a_j \\ \text{s.t. } a_j \geq 0 \\ \langle w, x_j \rangle - y_j \leq a_j \iff a_j - \langle w, x_j \rangle \geq -y_j \\ \langle w, x_j \rangle - y_j \geq -a_j \iff a_j + \langle w, x_j \rangle \geq y_j \end{cases}$$

Moreover, $\sum_j a_j = \langle u, c \rangle$ where c is the column vector in $\mathbb{R}^{d+m}$ with d zeros and m ones. Define a new matrix $X \in M_{m \times d}(\mathbb{R})$ by

$$X = \begin{pmatrix} - & x_1 & - \\ & \vdots & \\ - & x_m & - \end{pmatrix}$$

Observe now that the first constraint becomes

$$-\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \leq \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} - \begin{pmatrix} \sum_{j=1}^{d} w_j (x_1)_j \\ \vdots \\ \sum_{j=1}^{d} w_j (x_m)_j \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} - \begin{pmatrix} \sum_{j=1}^{d} w_j X_{1j} \\ \vdots \\ \sum_{j=1}^{d} w_j X_{mj} \end{pmatrix}$$

$$= \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} - Xw = \begin{pmatrix} -X & | & I_m \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_d \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} -X & | & I_m \end{pmatrix} u$$

Similarly, the second constraint becomes

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \leq \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} + Xw = \begin{pmatrix} X & | & I_m \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_d \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} X & | & I_m \end{pmatrix} u$$

Finally, define a new block matrix $A \in M_{2m \times (d+m)}$ and $b \in \mathbb{R}^{2m}$ by

$$A = \begin{pmatrix} -X & I_m \\ X & I_m \end{pmatrix}, \quad b = \begin{pmatrix} -y_1 \\ \vdots \\ y_m \\ y_1 \\ \vdots \\ y_m \end{pmatrix}$$

Then, in one expression, the above constraints become

$$Au = \begin{pmatrix} -X & I_m \\ X & I_m \end{pmatrix} u \geq \begin{pmatrix} -y_1 \\ \vdots \\ y_m \\ y_1 \\ \vdots \\ y_m \end{pmatrix} = b$$

Then, as linear program, our problem becomes

$$\begin{cases} \min_{u \in \mathbb{R}^{d+m}} \langle u, c \rangle \text{ such that} \\ Au \geq b \\ u \geq 0 \end{cases}$$

□

2 Problem 2

In this problem, we will get bounds on the VC-dimension of the class of (closed) balls in $\mathbb{R}^d$, that is,

$$\mathcal{B}_d = \{B_{v,r} : v \in \mathbb{R}^d, r > 0\}$$

where

$$B_{v,r}(x) = \begin{cases} 1 & \text{if } \|x - v\| \leq r \\ 0 & \text{otherwise.} \end{cases}$$

Consider the mapping $\varphi : \mathbb{R}^d \to \mathbb{R}^{d+1}$ defined by $\varphi(x) = (x, \|x\|^2)$. Show that if $x_1, \ldots, x_m$ are shattered by $\mathcal{B}_d$ then $\varphi(x_1), \ldots, \varphi(x_m)$ are shattered by the class of halfspaces in $\mathbb{R}^{d+1}$ (in this question we assume that $\operatorname{sign}(0) = 1$). What does this tell us about $\operatorname{VCdim}(\mathcal{B}_d)$?

Proof. Write $\mathcal{H} = \mathcal{B}_d$ and suppose that $S = \{x_1, \ldots, x_m\} \subseteq \mathbb{R}^d$ is shattered by $\mathcal{H}$, so $\{0,1\}^S = \mathcal{H}_S$. We wish to show that $(HS_{d+1})_S = \{0,1\}^{\varphi(S)}$. Suppose $f : \varphi(S) \to \{0,1\}$ is any function.

Let $g : S \to \{0,1\}$ be given by $g = f \circ \varphi$, so $g(x_i) = f(x_i, \|x_i\|^2)$. By shattering, there is some $v \in \mathbb{R}^d$ and $r > 0$ such that $g = B_{v,r}|_S$. We now have the following equivalences

$$\begin{aligned} f(\varphi(x_i)) = g(x_i) = 1 &\iff r \geq \|x_i - v\|^2 = \|x_i\|^2 + \|v\|^2 - 2\langle x_i, v\rangle \\ &\iff 0 \leq \langle 2v, x_i\rangle - \|x_i\|^2 - \|v\|^2 + r \\ &\iff 0 \leq \langle (2v, -1), \varphi(x_i)\rangle + r - \|v\|^2 = h_{(2v,-1),r-\|v\|^2}(\varphi(x_i)) \end{aligned}$$

Hence, we have that the half space $h_{(2v,-1),r-\|v\|^2}|_{\varphi(S)} = f$ as desired, whence $\varphi(S)$ is shattered by half spaces in $\mathbb{R}^{d+1}$. $\square$