Problem Set #2

Simas Glinskis

January 19, 2017

In collaboration with Natasha Antropova

Problem 1

We want to show that there is upper bound on $||\nabla_w l_{train}(w) - \nabla_w l_{gen}(w)||$, assuming that for all w, x, and y we have

$$||\nabla_w l(w, x, y)|| < b$$

First, let us compute the bounds on the gradients

$$||\nabla_w l_{train}(w)|| = ||\nabla_w \frac{1}{N} \sum_{i=1}^{N} l(w, x_i, y_i)|| = ||\frac{1}{N} \sum_{i=1}^{N} \nabla_w l(w, x_i, y_i)|| < \frac{1}{N} \sum_{i=1}^{N} b = b$$

and

$$||\nabla_w l_{gen}(w)|| = ||\nabla_w \sum_{i=1}^{N} p(x_i, y_i) l(w, x_i, y_i)|| = ||\sum_{i=1}^{N} p(x_i, y_i) \nabla_w l(w, x_i, y_i)|| < \sum_{i=1}^{N} p(x_i, y_i) b$$

Combining these,

$$||\nabla_w l_{train}(w) - \nabla_w l_{gen}(w)|| < b(1 - \sum_{i=1}^{N} p(x_i, y_i))$$

And this bound holds, and approaches 0 with infinite data.

Problem 2

See notebook with code and answers to parts a through d.