**Foundations of Deep Learning – Homework Assignment #4**Adi Album & Tomer Epshtein

**Part 2: (2)**

Question:  
Let be a hypothesis space corresponding to a neural network with parameters bounded in . For any subset , denote . Given a loss function and a training set , the Radamacher complexity of is defined to be:

Where:

* is short for

Assume that

Assume also that the implicit regularization of optimization leads to solutions with high , i.e., to:

Derive a generalization bound for that takes advantage of our knowledge on the implicit regularization, i.e. under which learned solutions with high ensure small generalization gap.

Proof:

Define for .  
Let be a small positive real value such that for some : . We define a series by:

* For all :
* For all

We have . For each , defines and this produces a series of subsets:

1. For every such that :   
    is a non-negative random variable, so by Markov’s Inequality:  
   :  
   We want to choose such that holds

I.e.  
Choose:  
So we have  
Or by viewing complement:

1. By proposition proved in class,  
    w.p. over :  
   Let , .
2. Reminder: For any two probabilistic events and we have

Let such that . Let’s look at:

* On the one hand:

* On the other hand:

Bringing it all together, we have:

such that , w.p.

Therefor, solutions with high are hypothesis’ with small , yielding small generalization gaps.