Generalization and Stability

Bousquet and Elisseeff
Presented by
Nilesh Kulkarni - 110050007
C. Yeshwanth - 110050083

November 22, 2014

Machine Learning in the SIL Setting

- The goal of a machine learning algorithm is to infer functions from data relating two sets of variables
- The input to a learning algorithm is a set of input-output pairs, $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ where $x_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$ where $\mathcal{X}$ and $\mathcal{Y}$
- The algorithm then outputs a function, $f : \mathcal{X} \to \mathcal{Y}$
- The hope is that the function f achieves low error on unseen data
- This quantity is estimated by a cost function $c : (\mathcal{Y}, \mathcal{Y}) \to \mathbb{R}$

What is Stability?

- The true distribution of the input variables is out of our hands
- The error is estimated based on the performance on the input(training) set
- A good learning algorithm should give similar results even if the input variables were changed slightly
- Stability answers questions of how changes in the input affect the output of the algorithm
- This notion of stability is then used to relate the performance of the function obtained by the algorithm on the training set to the performance on unseen data

Sources of Randomness

- There are two main sources of randomness in the input data
  - The first is due to the random distribution over the training data
  - The second is due to the noise encountered during the measurement of the variables
- This paper only considers the first kind of randomness
- The notion of stability used here is the notion of Uniform Stability
- The paper then goes on to show relationships between the stability of the learning algorithm and how well it generalizes
- It does not cover the question of statistical consistency

Uniform Stability

Definition

An algorithm A has uniform stability β with respect to loss function l if the following holds

$$\forall S \in \mathcal{Z}^m,\ \forall i \in \{1,\ldots,m\},\ \|\ell(A_S,.) - \ell(A_{S^{\setminus i}},.)\|_\infty \leq \beta$$

$l(f,z) = c(f(x),y)$ where $z = (x,y)$
A_S is the function returned by training set S
$A_{S^{\setminus i}}$ is the function returned by training set S with the i^{th} training example removed
A_{S^i} is the function returned by training set S with the i^{th} training example removed and replaced by a new training instance z'^i

Uniform Stability

- The condition imposed by this definition is that the performance on **any** new data point is not affected severely if one of the training points is removed
- We will show bounds on generalizablity using this notion of stability and some common algorithms which are uniformly stable

Bounds for Algorithms Exhibiting Uniform Stability

- If an algorithm exhibits uniform stability with parameter, β, a guarantee can be made about its generalization (Proof through McDiarmid)
- The Proof proceeds with finding a bound on the average error which is less than 2β
- Then it tries to find the maximum deviation from change in one data point $4\beta + \frac{M}{m}$
- Apply McDiarmid with $c_i's$ to get the bound
- β should decrease with at-least $O(\frac{1}{\sqrt{m}}))$ to achieve get meaningful results so that we can generalize

Stability of Regularization Algorithms

- We consider a class of loss function's which are σ -admissible
- The author introduces to *Bregman Divergence* to get some bounds on length of Δ f
- Bounds on Δ f can be used to derive bounds on the stability of the function returned by the learning algorithm
- Bounds β are proportional to length of $\phi(x)$ and inversely proportional to m

Conclusion

- The author establishes results which relate notions of stability to the property of generalization of learning algorithms
- These bounds were shown for a commonly used class of Hypothesis functions i.e. the RKHS with certain restrictions on the loss classes
- This analysis allows the analysis of learning algorithms where statistical consistency is not guaranteed
- This allows the user to trade off Statistical Consistency for greater generalization

Thank You
Questions?