Chapter 9 Review Notes

1 Stability of Machine Learning

1.1 Formal characterization

Let us assume that S is the dataset $S = z_1, z_1...z_i, ..z_n$

1.1.1 Version 1

$$\Delta_1(w) = ||w^*(S \cup K) - w^*(S)|| \text{ is bounded} \tag{1}$$

$$\Delta_1(w) = O(\frac{w_{max} + ...}{\lambda |S|}) \tag{2}$$

1.1.2 Version 2

Let us assume that $S_i^{'}$ is the dataset $S = z_1, z_1...z_i^{'}, ..z_n$

$$\Delta_2(w) = ||w^*(S) - w^*(S_i^{'})|| \text{ is bounded for all choice of } S \text{ and } S_i^{'} \tag{3}$$

$$\Delta_2(w) = O(\frac{2B}{\lambda |S|}) \tag{4}$$

where $B = \left[\frac{\partial l}{\partial w}\right]$ max element wise
$l(w, z)$ is point wise loss and the object was

$$F_s(w) = \sum_{i \in S} \left[\lambda ||w||^2 + l(w, z_i)\right] \tag{5}$$

1.2 Train, Validation and Test set

Dataset is divided into 3 parts(Train, Validation, Test).
Let us assume that S is the dataset and λ is the parameter.
In Train set we fix the λ and S and train for w. Suppose from the train set we get $w(\lambda, S)$.Now in Validation set we change the λ and measure the performance by putting λ in $w(\lambda, S)$.
We should note that test set is never seen practically. But we have to mimic the test set in some way hence validation set exist. Underlying assumption is that validation and test set should not be too much different.
Let us assume that $K(x_1, x_2)$ is our kernel function. So for fix K and λ we train w. Now assume that $L_s(w)$ is true loss and $F_s(w)$ is regularized loss

$$F_s(w) = \sum_{i \in S} \left[\lambda ||w||^2 + l(w, z_i)\right] \tag{6}$$

$$L_s(w) = \frac{\sum_{i \in S} l(w, z_i)}{|S|} \tag{7}$$

Let the Train set be S and Test set be D. Then the error in test set is

$$w^*(S) = F_s(w) \tag{8}$$

$$L_D(w) = \frac{\sum_{j \in D} l(w, z_j)}{|D|} \tag{9}$$

$$L_D(w^*(S)) = \text{ Test error after training w by minimizing } F_s(w) \tag{10}$$

$$L_D(w^*(S)) = L_s(w^*(S)) + [L_D(w^*(S)) - L_s(w^*(S))] \tag{11}$$

Avg-Test-error=Avg-Training-error+(Avg-Test-error - Avg-Training-error)

1.3 Max Training error

$$w^*(s) = argmin_w \frac{1}{|S|} \left[\sum_{i \in S} \lambda ||w||^2 + l(w, z_i) \right] \tag{12}$$

$$w^*(s) \leq \frac{1}{|S|} \sum_{i \in S} (0, z_i) \tag{13}$$

$$L_s(w) = \frac{\sum_{i \in S} l(w, z_i)}{|S|} \leq \left[\lambda ||w||^2 + \frac{\sum_{i \in S} l(w, z_i)}{|S|} \right] \leq \frac{1}{|S|} \sum_{i \in S} (0, z_i) \tag{14}$$

1.4 Max Test error

As we can see from above equations that our Test error is dependent on training set. We want our Test error to be independent of training set.
Let U be the universe of dataset.
Let $S \sim Uniform[U]$
Let N be the no of Samples of S To make test error independent of training set we can take expectation of test error over all samples of training set.

$$\mathbb{E}_s[L_D(w^*(S))] = \frac{1}{N} \sum_{k=1}^{N} L_D(w^*(S_k)) \tag{15}$$

$$\mathbb{E}_s[L_D(w^*(S))] = \mathbb{E}_s[L_s(w^*(S))] + (\mathbb{E}_s[L_D(w^*(S))] - \mathbb{E}_s[L_s(w^*(S))]) \tag{16}$$

Avg-Test-error=Avg-Training-error+(Avg-Test-error - Avg-Training-error)// Let us assume that $S = z_1, z_2..z_i...z_n$ and $S^{'} = z_1, z_2..z_i^{'}...z_n$

$$\text{Avg-Test-error - Avg-Training-error} = \mathbb{E}_s[L_D(w^*(S)) - L_s(w^*(S))]] \tag{17}$$

$$= \mathbb{E}_{s'}[L_D(w^*(S^{'})) - L_s(w^*(S))]] \tag{18}$$

$$= \mathbb{E}_{z \sim U}[l(w^*(S^{'}), z) - L(w^*(S), z)] \tag{19}$$

Eq 19 assumes that distribution of data in training and test set are same

$$\text{Test-error } \leq \text{Training-error} + \frac{K}{\lambda |S|} \tag{20}$$