Machine Learning – Assignment 1

103062528 林玉山

**1. 為什麼objective要用平方而不用絕對值？**

首先，這兩者共有的優點就是，得到的值都是正數，可以避免掉加總之後互相抵消的情況。那使用平方法的好處又在哪裡呢？

這個Objective的意義就是「預測的結果偏離實際結果的程度」，因此我們的目標就是最小化這個偏離程度。在數學上，最直覺的做法就是使用取一次導數 (derivative)，然後找到會使導數等於0的參數。如果我們採用了絕對值的objective，因為不具有連續性，所以在導數的處理上就會比平方法困難許多。

另外，二次式與絕對值的objective雖然都是convex function，但是二次式在接近optimal value時，會比絕對值的平滑許多。因此在使用gradient descent時，我們不需要找到optimal的數值，只要再該值附近，也能夠得到跟optimal差不多的結果。如果是絕對值，就需要進行比較多次的iteration才能夠逼近該值。

**2. 試證明**

因為 在training 過程中是定值，所以 正比於下列式子

代入

**3. Convex系列問題**

**(1) 試證明Convex Set的交集也是Convex Set**

假設有兩個Convex Set分別為A, B，A與B的交集則為Set C

我們從C中任取兩點x, y，因為C是A與B的交集，所以x, y也在A, B之中

因為A是Convex Set，所以根據定義，對於任意一個介於0到1之間的 ， 也會在A內；同時B也是Convex Set，對於任意一個介於0到1之間的 ， 也會在B內

因此， 同時存在於A與B內，代表也會存在於A, B的交集，C裡面

所以對於C內的任意點x, y， 也都會屬於C

因此C是Convex Set

**(2) 試證明上一題的 是Concave**

因為 是linear function，而 是convex function，而convex function的sum仍是convex function

因此 是convex function

接著，根據不同的t，得到的式子也仍是convex function，因此整體加起來之後仍會是convex function

而其負數 即為concave。

**4. Locally Weighted Linear Regression**

(a) 由該objective平方內的式子，是每一個sample x帶入w計算後與label r的差值，形式與 類似，因此可以猜測X應該就是每一份sample組成的matrix，r應該就是label組成的vector

而為了能夠使L分別與前後 相乘，所以須為n x n的matrix，內容應該就是weight

(b) 解Close Form

Linear Regression時，我們希望找出下式的w解

我們在左右式都乘上L

(c) 首先， 在training過程為定值，因此

又 ，所以

因此比較之後， 應為

(d)~(f) 請見README.pdf