Fengbo Ma

IE 7300 Section 1

Nov 16, 2022

Homework 3 – SMO Paper Summary

1. Overview

Sequential Minimal Optimization (known as SMO) is an algorithm for solving quadratic programming problem. The algorithm is invented by Platt, 1998. Not long after the invention of Support Vector Machine (SVM). For this course, we understand the algorithm by applying it to solve Support Vector Machine, especially the soft Margin Support Vector Machine.

1. Recap of Support Vector Machines

It is a common knowledge that the Soft Margin Support Vector Machine was being built based on the Vanilla SVM as shown below,

|  |  |  |
| --- | --- | --- |
|  |  | (1) |
|  |  | (2) |
|  |  | (3) |

Notice that, hard margin SVM no not work on non-linear boundary line nor overlap clusters. The problem is in its original format, where maximizing is not a traditional way of solving a convex optimization. The first step to the next for would be take the duel of the problem and make it into a minimizing problem. By introducing a slack variable to Hard Margin SVM, Soft Margin Support Vector Machine can be formed. In this report, we use variable σ. Hence, we have,

|  |  |  |
| --- | --- | --- |
|  |  | (3) |
|  |  | (4) |
|  |  | (5) |

With the slackness variation, total of four scenarios indicating four possible outcomes to indicate all the points on both side of the boundary line.

1. The Kernel Trick

As we could imagine, in both objective functions, function (1) and (3), the explicit expansion brought by create huge computational demand for only computing this part. It is the core reason that we most time heard that Support Vector Machine is not good for “big data” that would contain millions of records.

Rather than computing the explicit expansion that would bring the equation to the upper dimension. A kernel function in place to replace the explicit expansion could limiting some of computational time when applying the algorithm. The objective function now being updated to

|  |  |  |
| --- | --- | --- |
|  |  | (6) |
|  |  | (7) |

1. The Simplified SMO Algorithm

The Simplified SMO Algorithm even brought the algorithm to the next level, by not only applying Kernel Trick functions to it, it is also updating namely, a pair of α. By doing so, we are having a much rapidly when through down all the record. The updated Objective functions are shown as below,

|  |  |
| --- | --- |
|  | (8) |

To optimizing , we compute constrains that would bound the value of . We find the upper bound and lower bound (indicate by L and H) by equations shown below,

|  |  |
| --- | --- |
| : | (9) |
| : | (10) |

Now we could reach to an optimal and being presented by the function below,

|  |  |
| --- | --- |
|  | (11) |

Where E and η are being defined as the function given below. E are being defined as the error of SVM output,

|  |  |
| --- | --- |
|  | (12) |
|  | (13) |

By computing using functions above, can be computed accordingly with function here,

|  |  |
| --- | --- |
|  | (14) |

To handle the b threshold that KKT condition check brought us along all the examples,

|  |  |
| --- | --- |
|  | (15) |

And and are being defined as,

|  |  |  |
| --- | --- | --- |
|  | (16) | |
|  | (17) |