Machine Learning Optimization Model Using Green's Theorem

Farzad Gerami
Department of Industrial Engineering, Naghshejahan Higher Education Institute, Esfahan, 8144167984, Iran

Abstract

Optimization in the field of machine learning is one of the issues that experts are trying to use different approaches to provide a way to improve the machine performance. However, some mathematical tools mentioned in other sciences are very valuable to current issues in engineering science, who can enter and illuminate the field of engineering. In this paper, after a brief description of the issue raised by Green in microeconomic theory, tries to use from concept of "Separating Hyper plane" and "Supporting Hyper plane" to optimize the Support Vector Machine (SVM). Finally, a mathematical method is presented for use in such cases to maximizing margins in SVMs.

Keywords: Machine learning, Separating Hyper plane, Supporting Hyper plane.
1- Introduction
Due to the special place of the machine in different industries and its profound impact on the productivity of an enterprise, in recent years theorists have addressed issues related to the field of machine learning. Because of the importance of the issue, part of these researches is related to optimization. Olkopf and Smola (2001) with linear and nonlinear kernel and Shalev and Srebro (2008) with a reverse dependency approach, optimized SVM's. Also Xu et al. (2001) offered a method for clustering margin maximization. [1], [2], [3] Yang et al. (2016) optimized the SVM by a global stochastic optimization technique, particle swarm optimization (PSO) algorithm, it makes VW-SVM to be an adaptive parameter-free method for automated unmixing of protein subcellular patterns. [4] Aich and Banerjee (2016) inverse solution procedure is elaborated to find the near-optimum setting of process parameters in EDM machine to obtain the specific need based MRR-ASR combination. [5] Linn et al. (2016) proposed approach in the context of group classification using structural MRI data and showed that control-based normalization leads to better reproducibility of estimated multivariate disease patterns and improves the classifier performance in many cases. [6] In Ebrahimi and Khamenehi’s (2016) research support vector machine (SVM) was used to overcome the problem. The reservoir simulation software was replaced by the trained SVM. [7] In this study, using Green's theorem in the theory of microeconomics, a method have been developed to maximize margins in SVMs.

2- Green's theorem in the theory of microeconomics
In a matter of maximizing profit (UMP\(^2\)), the profit function \(U(x,y)\) is maximized due to some constraints such as budget \((1 - P_X \times X - P_Y \times Y = 0)\). Also, in a cost minimization problem (EMP\(^3\)), the cost function \(E(x,y)\) is minimized due to some constraints such as profit, where the total cost is \(E = P_X \times X + P_Y \times Y\). Using Lagrange multipliers, optimal vector in the UMP is synchronized with optimal integration in the EMP. Or in other words UMP and EMP are the dual problem. On the other hand Hessian matrix should be quite negative for a maximization problem and quite positive for a minimization problem. Mass-Colell et al. (1995) presented a different concept of duality. [8] In the mathematical appendix of this book, Green suggests the following two important cases:
   a. Separating Hyper plane Theorem
   b. Supporting Hyper plane Theorem
The following are briefly discussed.

---
\(^1\) Support Vector Machine
\(^2\) Utility Maximization Problem
\(^3\) Expenditure Minimization Problem
2-1- Separating Hyper plane Theorem

Assume $A, B \subseteq R^n$ and $A \cap B = \emptyset$. Also assume collection $B$ is a convex and closed collection as $x \in B$ and $y \in B$. There is $p \in R^n$, $p \neq 0$ per $x, y$ and there is $c \in R^n$ as $p \cdot x > c, p \cdot y < c$. There is therefore a top plate ($H_{p,c}$) that separates the A and B, that they (A and B) are opposite each other on either side.

2-2- Supporting Hyper plane Theorem
Assume $B \subseteq R^n$ is a convex and $x \in B$. There is $p \in R^n$, $p \neq 0$ per $y$ as $p.x \geq p.y$. There is therefore a top plate ($H_{p,c}$) that support $B$.

3- Optimization mathematical model in SVMs
According to the two cases and Green’s approach in microeconomic, we can suggest a method to optimize the SVMs. The use of this theorem is that, based on a mathematical method, a SVM should be able to choose the best plate. If we have the following definition:
Separator plate: $x.w + b = 0$
Supporting plate: $x.w + b = -1$
Margin: $1/||w||$
Normal vector: $w$

Then, based on this mathematical method, SVMs select the plate that maximize the margin of separation between the two classes from all the separator plates. A SVM classifies the inputs into two classes, using a plate in a multi-dimensional space. A SVM includes vector $b$ and vector $w$ that use them to training data $\{x, y\}$. It is described as follows:

$$x_i.w + b \geq +1 \text{ for } y_i = +1$$
and
$$x_i.w + b \leq -1 \text{ for } y_i = -1$$
or
$$y_i(x_i.w + b) - 1 \geq 0 \quad \forall i$$

Margin is defined using vector geometry $1/||w||$, where $w$ is the normal vector to a separator plate which is equidistant from the supporting plate, so that the supporting plates support various classes or series of observations in the training data.
Maximization of margins is an optimization problem, which involves the use of Lagrange multipliers, which is common in economics. In Lagrange settings, where $H_{ij} = y_i y_j x_i x_j$, the term $\alpha^T H \alpha$ is formulated. In the study, "Linear Kernel" is intended as a point production, $K(x_i, x_j) = x_i^T x_j$. If the data are not linearly separable, other kernel functions such as polynomial and annular kernel functions is recommended for non-linear separable data.

4- Conclusion

In this study, after two brief description of Green's theorem, was tried to use from mathematical concepts in the field of machine learning. The findings of this study show that Green's theorem in microeconomic theory can be used to maximize margins SVMs. This research could open the window to develop new approach to optimization of SVMs. Based on the results of this study, it is recommended for future research an algorithm be designed and implemented.
References