

Separating Hyperplane Problem

Statement

Given a set of data points labeled as positive or negative, find a separating hyperplane that maximizes the margin between the hyperplane and the closest data points. Formulate this problem as a convex optimization problem and explain how it relates to Support Vector Machines (SVMs) for classification.

Mathematical Formulation

Assume we have a dataset with n data points. Each data point is represented by a vector $\mathbf{x}_i = (x_1, x_2, \dots, x_n)$ and a corresponding label y_i that can be either $+1$ or -1 . Let the separating hyperplane be

$$w^T \mathbf{x} + b = 0 \quad (1)$$

where w is a normal vector to the hyperplane and b is the bias term.

****Important Note:**** While the formulation above shows the margin for a single data point, the objective is to find the hyperplane that maximizes the margin for all data points simultaneously.

The margin of a data point \mathbf{x}_i from the hyperplane is given by

$$\frac{|w^T \mathbf{x}_i + b|}{\|w\|} \quad (2)$$

In the context of multiple data points, we want to consider the minimum margin across all data points:

$$\min_{i=1}^n \frac{|w^T \mathbf{x}_i + b|}{\|w\|} \quad (3)$$

The goal is then to maximize this minimum margin.

We can choose w and b such that

$$\min_{i=1}^n |w^T \mathbf{x}_i + b| = 1 \quad (4)$$

The rest of the formulation follows the same logic, incorporating the concept of multiple data points with summations where applicable, to arrive at the equivalent convex optimization problem:

$$\text{minimize } \frac{1}{2} \|w\|^2 \text{ subject to } y_i(w^T \mathbf{x}_i + b) \geq 1 \text{ for all } i = 1, 2, \dots, n.$$

This optimization problem is convex.

Relation with SVM

The problem statement given, to find a separating hyperplane that maximizes the margin between the hyperplane and the closest data points, is the fundamental objective of Support Vector Machines (SVMs) for classification. SVMs aim to find the optimal hyperplane that separates data points of different classes while maximizing the margin between the hyperplane and the nearest data points from each class. This is achieved by formulating the problem as a convex optimization problem, exactly as described in the mathematical formulation section above.

Therefore, the problem statement described is directly related to the core objective of SVMs, making it a fundamental concept in SVM classification.