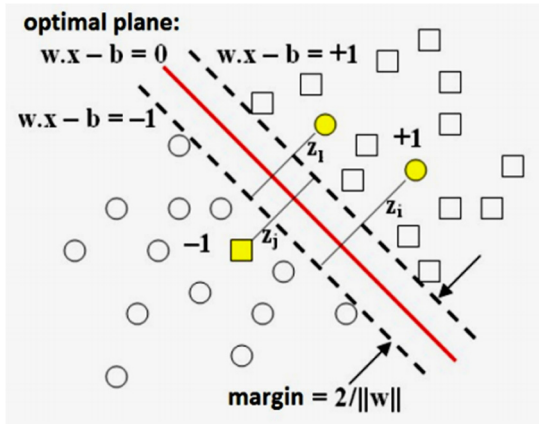


3.6 Máy học véc tơ hỗ trợ (Support Vector Machines - SVM)

Máy học véc tơ hỗ trợ (SVM) được đề xuất bởi (Vapnik, 1995) là mô hình hiệu quả và phổ biến cho vấn đề phân lớp, hồi quy những tập dữ liệu có số chiều lớn.

Xét ví dụ phân lớp nhị phân tuyến tính (hình 6) với m phần tử x_1, x_2, \dots, x_m trong không gian n chiều, có nhãn (lớp) của các phần tử là y_1, y_2, \dots, y_m có giá trị 1 hoặc -1 . $y_i = 1$, nếu x_i thuộc lớp $+1$ (lớp dương, lớp chúng ta quan tâm), $y_i = -1$, nếu x_i thuộc lớp -1 (lớp âm hay các lớp còn lại). SVM tìm siêu phẳng tối ưu (xác định bởi véc tơ pháp tuyến w và độ lệch của siêu phẳng b) dựa trên 2 siêu phẳng hỗ trợ của 2 lớp. Các phần tử lớp $+1$ nằm bên phải của siêu phẳng hỗ trợ cho lớp $+1$, các phần tử lớp -1 nằm phía bên trái của siêu phẳng hỗ trợ cho lớp -1 . Những phần tử nằm ngược phía với siêu phẳng hỗ trợ được coi như lỗi. Khoảng cách lỗi được biểu diễn bởi $z_i \geq 0$ (với x_i nằm đúng phía của siêu phẳng hỗ trợ của nó thì khoảng cách lỗi tương ứng $z_i = 0$, còn ngược lại thì $z_i > 0$ là khoảng cách từ điểm x_i đến siêu phẳng hỗ trợ tương ứng của nó).



Hình 6: Phân lớp tuyến tính với máy học véc tơ hỗ trợ

Khoảng cách giữa 2 siêu phẳng hỗ trợ được gọi là lề. Siêu phẳng tối ưu (nằm giữa 2 siêu phẳng hỗ trợ) tìm được từ 2 tiêu chí là cực đại hóa lề (lề càng lớn, mô hình phân lớp càng an toàn) và cực tiểu hóa lỗi. Vấn đề dẫn đến việc giải bài toán quy hoạch toàn phương (11):

$$\min \Psi(w, b, z) = (1/2) \|w\|^2 + c \sum_{i=1}^m z_i \quad s.t \quad (11)$$

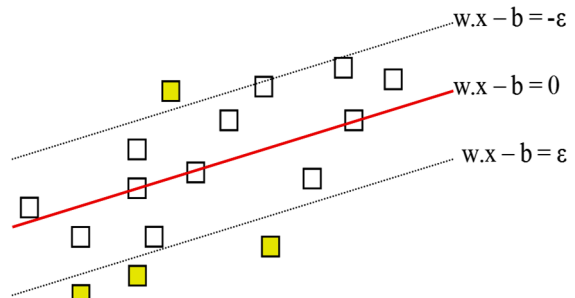
$$y_i(w \cdot x_i - b) + z_i \geq 1$$

$$z_i \geq 0 \quad (i=1, 2, \dots, m)$$

với hằng $c > 0$ được sử dụng để chỉnh độ rộng lề và lỗi.

Giải bài toán quy hoạch toàn phương (11), thu được (w, b) . Phân lớp phần tử x dựa vào dấu của $(w \cdot x - b)$ như trong (12). Nếu giá trị biểu thức $(w \cdot x - b) > 0$ thì gán nhãn cho x là lớp dương ($+1$), ngược lại thì gán nhãn cho x là lớp âm (-1).

$$\text{predict}(x) = \text{sign}(w \cdot x - b) \quad (12)$$



Hình 7: Hồi quy với máy học véc tơ hỗ trợ

Máy học SVM cũng có thể xử lý bài toán hồi quy. Trong vấn đề hồi quy như Hình 7, SVM tìm siêu phẳng đi qua tất cả các phần tử dữ liệu với độ lệch chuẩn là ϵ . Huấn luyện máy học SVM cho xử lý vấn đề hồi quy dẫn đến việc giải bài toán quy hoạch toàn phương (13) như sau:

$$\begin{aligned} \min \Psi(w, b, z^*, z) &= (1/2) \|w\|^2 + \\ c \sum_{i=1}^m (z_i^* + z_i) \quad s.t. \quad (13) \\ w \cdot x_i - b - y_i - z_i^* &\leq \epsilon \\ w \cdot x_i - b - y_i + z_i &\geq -\epsilon \\ z_i^*, z_i &\geq 0 \quad (i=1, 2, \dots, m) \end{aligned}$$

với hằng $c > 0$ được sử dụng để chỉnh độ rộng lề và lỗi.

Giải bài toán quy hoạch toàn phương (13) sẽ thu được siêu phẳng hồi quy (w, b) của SVM. Dự báo cho phần tử mới đến x dựa trên siêu phẳng (w, b) được tính theo công thức (14):

$$\text{predict}(x) = (w \cdot x - b) \quad (14)$$

Giải thuật SVM có thể thay thế các tích vô hướng trong các công thức (11-14) bởi hàm nhân (kernel functions), sẽ cho phép giải quyết một số lớn các bài toán phân lớp và hồi quy phi tuyến. Không có bất kỳ một thay đổi nào cần thiết về mặt