FH Vorarlberg
University of Applied Sciences

Support Vector Machines (SVM)

Computational Intelligence II

Informatik - Software and Information Engineering
Fachhochschule Vorarlberg

Erstellt von
André Hopfgartner & Matthias Rupp

Dornbirn, am 5. März 2021

Inhaltsverzeichnis

Abkürzungsverzeichnis

SVM Support Vector Machine

1 Einführung

1.1 Intuition

Ziel: möglichst breites Band zwischen den 2 verschiedenen Klassen aufziehen.

1.2 Mathematische Herleitung

TODO TEXT HERE

1.2.1 Problemdefinition

Gegeben sei ein Gewichtsvektor $w \in \mathbb{R}^D$, ein Bias $b \in \mathbb{R}$ und ein beliebiger Punkt $x \in \mathbb{R}^D$. Eine Ebene im Raum kann definiert werden durch:

$$w^T x + b = 0 \tag{1.1}$$

Weil Gleichung 1.1 mit verschiedenen Skalarwerten skaliert werden kann, führen wir eine zusätzliche Bedingung ohne Beschränkung der Allgemeinheit ein. Sei $x_n \in \mathbb{R}^D$ der am nächsten zur Ebene gelegene Punkt so soll gelten:

$$|w^T x_n + b| = 1 \tag{1.2}$$

Als nächsten Schritt bestimmen wir den euklidischen Normalabstand D eines beliebigen Punkts $x_k \in \mathbb{R}^D$ zu der Ebene. Hierfür ist zuerst zu bemerken, dass w normal zur definierten Ebene steht.

Lemma 1.2.1. *Eine Ebene sei definiert durch $w^T x + b = 0$. Der Vektor w steht normal zu der definierten Ebene.*

Beweis. Man wähle zwei Punkte $x_1, x_2 \in \mathbb{R}^D$ die auf der Ebene liegen. Somit muss gelten:

$$\begin{aligned} w^T x_1 + b &= 0 \\ w^T x_2 + b &= 0 \\ w^T (x_1 - x_2) = 0 \leftrightarrow \|w^T\| \|x_1 - x_2\| \cos(\alpha) &= 0 \leftrightarrow \alpha = 90^\circ \end{aligned} \tag{1.3}$$

□

Um den Normalabstand D eines beliebigen Punkts x_k zu ermitteln wählt man einen Punkt x der auf der Ebene liegt und projiziert den Vektor $(x_k - x)$ auf den

Einheitsvektor von w. Weil nur der tatsächliche Abstand zur Ebene relevant ist nimmt man den Betrag.

$$
\begin{aligned}
D &= |\frac{w^T}{\|w\|}(x_k - x)| = \\
&= \frac{1}{\|w\|}|(w^T x_k - w^T x)| = \\
&= \frac{1}{\|w\|}|(w^T x_k + b - (w^T x + b))|
\end{aligned}
\tag{1.4}
$$

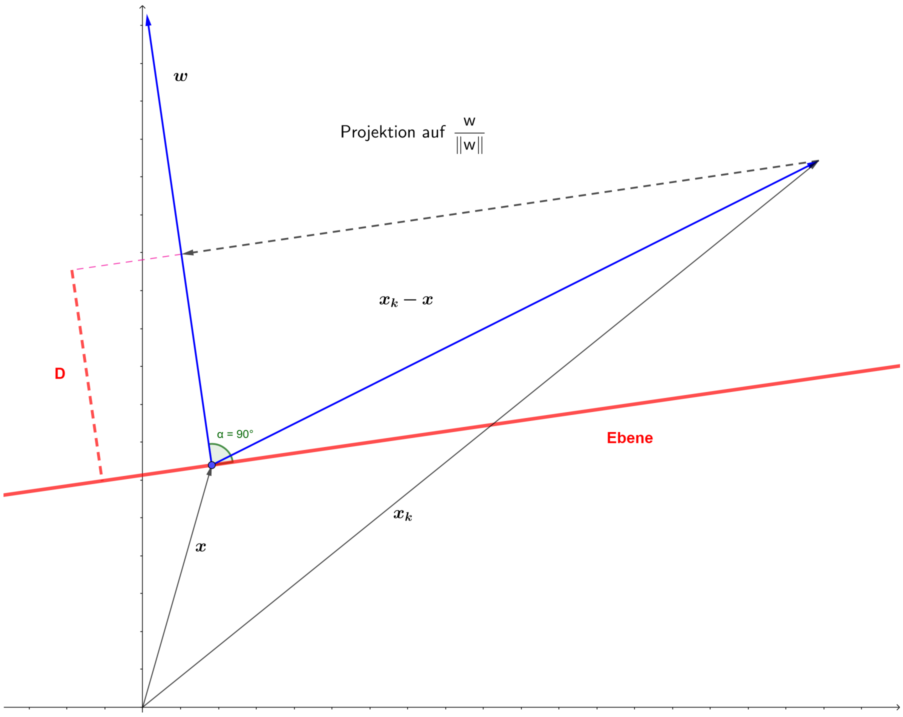

Abbildung 1.1: Durch die Projektion von $(x_k - x)$ auf den Einheitsvektor von w kann der Normalabstand D von x_k zu der Ebene bestimmt werden.

Weil der Punkt x auf der Ebene liegt gilt $w^T x + b = 0$ (Gleichung 1.1):

$$D = \frac{1}{\|w\|}|(w^T x_k + b)| \tag{1.5}$$

Aus Gleichung 1.2 gilt weiters $|(w^T x_k + b)| = 1$ wenn x_k der am nächsten zu der Ebene liegende Punkt ist. Somit ergibt sich der kleinste Abstand zur Ebene als:

$$D = \frac{1}{\|w\|} \tag{1.6}$$

1.2.2 Optimierungsproblem

Gleichung 1.6 beschreibt den Normalabstand zu dem am nächsten an der Ebene liegenden Punkt. Ziel einer Support Vector Machine (SVM) ist die Maximierung dieses Abstands. Hierbei handelt es sich um ein Optimierungsproblem mit Nebenbedingungen:

$$\max_{w} \quad \frac{1}{\|w\|} \tag{1.7a}$$

$$\text{mit} \quad \min_{n=1..N} |w^T x_n + b| = 1 \tag{1.7b}$$

Der Betrag lässt sich umschreiben durch die Einführung von Labels $y_n \in \{-1, +1\}$. Diese beschreiben jeweils die Klasse des zugehörigen Vektors x_n. Für eine korrekte Klassifizierung der SVM gilt:

$$y_n = sign(w^T x_n + b) \tag{1.8}$$

Somit gilt für einen korrekt klassifizierten Vektor x_k:

$$|w^T x_n + b| = y_n(w^T x_n + b) \tag{1.9}$$

Durch Anwendung von Gleichung 1.9 in Gleichung 1.7b, Umformulierung der zu optimierenden Funktion in eine Minimierung und der Verallgemeinerung auf beliebige Punkte x_n erhält man:

$$\min_{w} \quad \frac{1}{2} w^T w \tag{1.10a}$$

$$\text{mit} \quad y_n(w^T x_n + b) \geq 1 \text{ für } n = 1..N \tag{1.10b}$$

Die Verallgemeinerung von Gleichung 1.7b auf Gleichung 1.10b auf beliebige Punkte ist so möglich, weil durch Gleichung 1.2 sichergestellt ist, dass der kleinste Wert für $(w^T x_n + b)$ 1 ist, und somit die Werte für alle anderen Punkte größer oder gleich 1 sein müssen.