Neuronale Netze

Prof. Dr. Klaus Meer, M. Sc. Ameen Naif

Aufgabenblatt 4
14. Juni 2017

Aufgabe 1.

Betrachten Sie die Abbildung $\varphi : \mathbb{R} \to \mathbb{R}^{d+1}$ gemäß

$$\varphi(x) = (a_0 \cdot 1, a_1 \cdot x, a_2 \cdot x^2, \ldots, a_d \cdot x^d), a_i \in \mathbb{R} \ .$$

Die Kernfunktion K wird nun definiert durch $k(x, y) := \varphi(x)^T \cdot \varphi(y)$.

a) Berechnen Sie K für den Fall, dass alle a_i's als 1 gewählt werden.

b) Finden Sie geeignete Werte a_i so, dass man den Kern $k(x, y) = (1+x\cdot y)^d$ erhält. Vergleichen Sie Ihr Ergebnis mit dem Kern in a).

c) Zeigen Sie im Fall a) folgendes: für jede Wahl von $d + 1$ verschiedenen Punkten $x_i, 0 \leq i \leq d$ und für jedes beliebige 'Vorzeichenmuster' $d_i \in \{-1, 1\}, 0 \leq i \leq d$ gibt es ein eindeutiges $w \in \mathbb{R}^{d+1}$ mit $w^t \cdot \varphi(x_i) = d_i$ für alle i.

Aufgabe 2.

Betrachten Sie einen Eingaberaum der Dimension n. Definieren Sie einen Kern $k(x, y)$ über

$$k(x, y) = (1 + x \cdot y)^d$$

für $x, y \in \mathbb{R}^n, d \in \mathbb{N}$ fest.

a) Zeigen Sie, dass man für ein geeignetes endliches N eine Feature-Abbildung $\varphi : \mathbb{R}^n \to \mathbb{R}^N$ mit $k(x, y) = \varphi(x)^T \cdot \varphi(y)$ definieren kann. Sie sollen nur die Existenz von φ zeigen, berechnen müssen Sie es nicht.

b) Berechnen Sie für $d := 3, n := 2$ in a) die Abbildung φ.

Aufgabe 3.

Wir betrachten erneut das XOR-Problem, diesmal mit $C_1 = \{(1, 1)^T, (-1, -1)^T\}$ und $C_2 = \{(1, -1)^T, (-1, 1)^T\}$. Punkte in C_1 erhalten als Zielwert -1, die in C_2 den Wert 1.

Als Kernfunktion wählen wir $k(x, y) := (x^T \cdot y + 1)^2$.

a) Stellen Sie das zugehörige duale Problem auf. Bestimmen Sie dazu zunächst die zugehörigen Feature-Abbildungen (6 Stück).

b) Lösen Sie das Problem in a). Einer der 4 Lagrangeparameter lässt sich unmittelbar eliminieren. überlegen Sie für die verbleibenden Parameter, ob sie zu aktiven Nebenbedingungen gehören.

c) Stellen Sie die trennenden Hyperebene im Featureraum (also hier $\mathbb{R}^6$) dar.

Aufgabe 4.

In dieser Aufgabe soll untersucht werden, welche Form Kernfunktionen haben, die für endliche Testmengen definiert sind. Sei dazu $X = \{x^1, \ldots, x^\ell\} \subset \mathbb{R}^n$ endlich. Eine Funktion $K : X \times X \to \mathbb{R}$ heißt Kernfunktion auf X, sofern es ein geeignetes $m \in \mathbb{N}$ und $\varphi : X \to \mathbb{R}$ gibt mit $K(x, y) = \varphi(x)^T \cdot \varphi(y)$ (Skalarprodukt in $\mathbb{R}^m$).

Definiere für eine beliebige Funktion $K : X \times X \to \mathbb{R}$ eine $\ell \times \ell$ Matrix $\mathcal{K}$ über $\mathcal{K}_{ij} := K(x^i, x^j)$.

Zeigen Sie:

a) Falls K eine Kernfunktion ist, dann hat die symmetrische Matrix $\mathcal{K}$ nur nicht-negative Eigenwerte.

b) Ist umgekehrt $\mathcal{K}$ positiv semi-definit, dann *ist K eine Kernfunktion.*

Aufgabe 5.

Sei $k(x, y) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine durch $k(x, y) = \varphi(x)^T \cdot \varphi(y)$ gegebene Kernfunktion, wobei $\varphi : \mathbb{R}^n \times \mathbb{R}^m$ mit m fest gegeben sei.

a) Zeigen Sie für alle $x, y \in \mathbb{R}^n : k(x, y) \leq k(x, x) \cdot k(y, y)$

b) Definieren für festes $x \in \mathbb{R}^n$ die Funktion

$$k(\bullet, x) : \mathbb{R}^n \to \mathbb{R} \ , \ y \to k(y, x) \ .$$

Sei ferner $\mathcal{F}$ der lineare Raum aller Funktionen der Gestalt

$$f(x) := \sum_{i=1}^{\ell} s_i \cdot k(x, x_i)$$

mit $\ell \in \mathbb{N}, x_i \in \mathbb{R}^n, s_i \in \mathbb{R}$ beliebig (d.h. für verschiedene Funktionen in $\mathcal{F}$ dürfen diese Werte verschieden gewählt werden).

In $\mathcal{F}$ definieren wir ein Skalarprodukt vermittels

$$\langle f, g \rangle_{\mathcal{F}} := \sum_{i=1}^{\ell} \sum_{j=1}^{k} s_i t_j \cdot k(x_i, x'_j)$$

mit $f(x) := \sum_{i=1}^{\ell} s_i \cdot k(x, x_i), \quad g(x) := \sum_{j=1}^{k} t_j \cdot k(x, x'_j)$.

Zeigen Sie:

i) $\langle k(\bullet, x), f \rangle_{\mathcal{F}} = f(x) \ \forall x \in \mathbb{R}^n$

ii) $\langle k(\bullet, x), k(\bullet, x') \rangle_{\mathcal{F}} = k(x, x') \ \forall x, x' \in \mathbb{R}^n$

iii) $\langle f, f \rangle_{\mathcal{F}} = 0 \ \Leftrightarrow \ f \equiv 0$.

Aufgabe 6.

Ist zu einer gegebenen Kernfunktion k die Feature-Abbildung φ mit $k(x, y) = \varphi(x)^T \cdot \varphi(y)$ eindeutig bestimmt? Ist die Diemnsion des Bildraums eines zu k passenden φ eindeutig bestimmt?

Aufgabe 7.

Gegeben seien die beiden Punktmengen: $C_1 = \{(1,0)^T, (2,1)^T\}$ und $C_2 = \{(-2,0)^T, (-2,-1)^T\}$ im $\mathbb{R}^2$. Die Punkte werden der angegebenen Reihenfolge nach mit $x_1, \ldots, x_4$ bezeichnet, die Zielwerte d_i seien dementsprechend 1 für Punkte aus C_1 und -1 für Punkte aus C_2.

Wie lauten für eine lineare SVM (d.h. ohne Fehlerterm zur Korrektur verletzter Nebenbedingungen) das primale und das duale Optimierungsproblem?

Aufgabe 8.

Gegeben seien die beiden Punktmengen
$C_1 = \{(1,0)^T, (2,0)^T, (3,0)^T, (1,1)^T, (2,1)^T, (2,3)^T, (4,1)^T\}$ und
$C_2 = \{(-1,0)^T, (-1,1)^T, (-2,0)^T, (-2,-1)^T\}$ im $\mathbb{R}^2$.

Argumentieren Sie geomterisch, welche Punkte bei einer linearen SVM Support-Vektoren werden können. Was ist die optimale Trennebene?

Aufgabe 9.

Welche der folgenden Funktionen $k : \mathbb{R}^n \to \mathbb{R}$ ist eine Kernfunktion? Begründen Sie Ihre Antworten.

a) $k(x,y) = (x^T \cdot y + 1)^3$

b) $k(x,y) = -(x^T \cdot y + 1)^3$

c) $k(x,y) = (x^T \cdot x)^4 + (y^T \cdot y)^2$.