

Maschinelles Lernen und Data Mining
Sommersemester 2014
Übungsblatt 4

Besprechung des Übungsblattes am 22.05.2014

Aufgabe 4-1 Kernkombinationen

Um einen selbstdefinierten Kern $k(\mathbf{x}_i, \mathbf{x}_j)$ für $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^n$ anzuwenden, muss gezeigt werden, dass es sich auch tatsächlich um einen legitimen Kern handelt. Da es recht aufwendig sein kann zu zeigen, dass für k das *Mercer Theorem* zutrifft, wird häufig explizit das Mapping der impliziten Basistransformationen angegeben: $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$.

Eine weitere beliebte Variante, die Gültigkeit einer Kernfunktion zu zeigen, ist die Rückführung auf eine Kombination aus Kernen, da für einige Operationen $\circ$ gilt, dass $k(\mathbf{x}_i, \mathbf{x}_j) = k_1(\mathbf{x}_i, \mathbf{x}_j) \circ k_2(\mathbf{x}_i, \mathbf{x}_j)$ ein legitimer Kern ist.

Zeigen Sie, dass für valide Kernfunktionen $k_l(\mathbf{x}_i, \mathbf{x}_j)$ mit $l \in \mathbb{N}_+$ gilt:

- a) **Skalierung:** Für $a > 0$ ist $k(\mathbf{x}_i, \mathbf{x}_j) := a \cdot k_1(\mathbf{x}_i, \mathbf{x}_j)$ ein Kern.
- b) **Summe:** $k(\mathbf{x}_i, \mathbf{x}_j) := k_1(\mathbf{x}_i, \mathbf{x}_j) + k_2(\mathbf{x}_i, \mathbf{x}_j)$ ist ein Kern.
- c) **Linearkombination:** Für $w \in \mathbb{R}_+^d$ ist $k(\mathbf{x}_i, \mathbf{x}_j) := \sum_{l=1}^d w_l \cdot k_l(\mathbf{x}_i, \mathbf{x}_j)$ ein Kern.
- d) **Produkt:** $k(\mathbf{x}_i, \mathbf{x}_j) := k_1(\mathbf{x}_i, \mathbf{x}_j) \cdot k_2(\mathbf{x}_i, \mathbf{x}_j)$ ist ein Kern.
- e) **Potenz:** Für ein $p \in \mathbb{N}_+$: $k(\mathbf{x}_i, \mathbf{x}_j) := (k_1(\mathbf{x}_i, \mathbf{x}_j))^p$ ist ein Kern.

Lösungsvorschlag:

a)

$$\begin{aligned} ak_1(\mathbf{x}_i, \mathbf{x}_j) &= a\phi_1(\mathbf{x}_i)^T \phi_1(\mathbf{x}_j) = a \sum_{m=1}^M \phi_1(\mathbf{x}_i)_m \phi_1(\mathbf{x}_j)_m = \sum_{m=1}^M \sqrt{a}\phi_1(\mathbf{x}_i)_m \sqrt{a}\phi_1(\mathbf{x}_j)_m = \\ &= \sum_{m=1}^M \phi'_1(\mathbf{x}_i)_m \phi'_1(\mathbf{x}_j)_m = \phi'_1(\mathbf{x}_i)^T \phi'_1(\mathbf{x}_j) \end{aligned}$$

mit $\phi'_1(\mathbf{x}_i) = \sqrt{a}\phi_1(\mathbf{x}_i)$

b)

$$k_1(\mathbf{x}_i, \mathbf{x}_j) + k_2(\mathbf{x}_i, \mathbf{x}_j) = \phi_1(\mathbf{x}_i)^T \phi_1(\mathbf{x}_j) + \phi_2(\mathbf{x}_i)^T \phi_2(\mathbf{x}_j) = \phi''_{12}(\mathbf{x}_i)^T \phi''_{12}(\mathbf{x}_j)$$

mit $\phi''_{12}(\mathbf{x}_i) = \begin{pmatrix} \phi_1(\mathbf{x}_i) \\ \phi_2(\mathbf{x}_i) \end{pmatrix}$

c)

$$\sum_{l=1}^n w_l k_l(\mathbf{x}_i, \mathbf{x}_j) \stackrel{\text{a)}}{=} \sum_{l=1}^n \phi'_l(\mathbf{x}_i) \phi'_l(\mathbf{x}_j) \stackrel{\text{b)}}{=} \phi''_{1\dots l}(\mathbf{x}_i)^T \phi''_{1\dots l}(\mathbf{x}_j)$$

mit $\phi'_l(\mathbf{x}_i) = \sqrt{w_l}\phi_l(\mathbf{x}_i)$ und $\phi''_{1\dots l}(\mathbf{x}_i) = \begin{pmatrix} \phi'_1(\mathbf{x}_i) \\ \phi'_2(\mathbf{x}_i) \\ \vdots \\ \phi'_l(\mathbf{x}_i) \end{pmatrix}$

b.w.

Lösungsvorschlag:

d)

$$\begin{aligned}
 k_1(\mathbf{x}_i, \mathbf{x}_j) \cdot k_2(\mathbf{x}_i, \mathbf{x}_j) &= \left(\sum_{m=1}^{M_1} \phi_1(\mathbf{x}_i)_m \cdot \phi_1(\mathbf{x}_j)_m \right) \cdot \left(\sum_{n=1}^{M_2} \phi_2(\mathbf{x}_i)_n \cdot \phi_2(\mathbf{x}_j)_n \right) = \\
 &= \sum_{m=1}^{M_1} \left(\phi_1(\mathbf{x}_i)_m \phi_1(\mathbf{x}_j)_m \cdot \left(\sum_{n=1}^{M_2} \phi_2(\mathbf{x}_i)_n \phi_2(\mathbf{x}_j)_n \right) \right) = \\
 &= \sum_{m=1}^{M_1} \sum_{n=1}^{M_2} (\phi_1(\mathbf{x}_i)_m \phi_1(\mathbf{x}_j)_m \cdot \phi_2(\mathbf{x}_i)_n \phi_2(\mathbf{x}_j)_n) = \\
 &= \sum_{m=1}^{M_1} \sum_{n=1}^{M_2} (\phi_1(\mathbf{x}_i)_m \phi_2(\mathbf{x}_i)_n) \cdot (\phi_1(\mathbf{x}_j)_m \phi_2(\mathbf{x}_j)_n) = \\
 &= \phi_{12}'''(\mathbf{x}_i)^T \phi_{12}'''(\mathbf{x}_j) = k(x_i, x_j)
 \end{aligned}$$

$$\text{mit } \phi_{12}'''(\mathbf{x}_i) = \begin{pmatrix} \phi_1(\mathbf{x}_i)_1 \cdot \phi_2(\mathbf{x}_i)_1 \\ \phi_1(\mathbf{x}_i)_1 \cdot \phi_2(\mathbf{x}_i)_2 \\ \vdots \\ \phi_1(\mathbf{x}_i)_1 \cdot \phi_2(\mathbf{x}_i)_{M_2} \\ \phi_1(\mathbf{x}_i)_2 \cdot \phi_2(\mathbf{x}_i)_1 \\ \vdots \\ \phi_1(\mathbf{x}_i)_{M_1} \cdot \phi_2(\mathbf{x}_i)_{M_2} \end{pmatrix}$$

e) Beweis per Induktion:

I.V.: $k(\mathbf{x}_i, \mathbf{x}_j) := k_1(\mathbf{x}_i, \mathbf{x}_j)^p$ ist ein Kern (für p positives Integer).

I.A.: $k_1(x_i, x_j) \cdot k_1(x_i, x_j)$ ist ein legaler Kern. Gezeigt in d).

I.S.: $k_1(\mathbf{x}_i, \mathbf{x}_j)^{p+1} = k_1(\mathbf{x}_i, \mathbf{x}_j)^p \cdot k_1(\mathbf{x}_i, \mathbf{x}_j) \stackrel{\text{I.V.}}{=} \text{Kern} \cdot k_1(\mathbf{x}_i, \mathbf{x}_j) \stackrel{\text{d)}}{=} \text{valider Kern.}$