Ludwig-Maximilians-Universität München
Institut für Informatik
PD Dr. Peer Kröger, Dr. Arthur Zimek

Knowledge Discovery in Databases
WS 2010/11

Übungsblatt 7: Kernels und Support-Vektor-Machines

Lösung zu 7-2

(a) Optimierungsproblem mittels Lagrange:

Achtung: y bezeichnet nicht die Koordinate, sondern die Klasse, d.h. $+1$ bzw. -1.

Gegeben: $\vec{x_1} = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$ $y_1 = -1$ $\vec{x_2} = \begin{pmatrix} 2 & 2 \end{pmatrix}^T$ $y_2 = +1$

Wir haben zwei Support Vectors, daher gibt es α_1 und α_2. Da aber $\sum_i \alpha_i y_i = 0 = \alpha_1 y_1 + \alpha_2 y_2 = \alpha_2 - \alpha_1$ gilt $\alpha_1 = \alpha_2 =: \alpha$.

Die Formel aus der Vorlesung in Kernel-Schreibweise lautet:

$$L(\vec{\alpha}) := \left(\sum_i \alpha_i\right) - \frac{1}{2}\left(\vec{\alpha}^T M \vec{\alpha}\right)$$

Diese Kernel-Schreibweise erlaubt die Verwendung von beliebigen Kerneln!

Mit $M_{ij} := y_i y_j \langle \vec{x_i}, \vec{x_j} \rangle$ (die Summenschreibweise erhält man durch Ausmultiplizieren!)

Wir berechnen die Kernelmatrix und setzen sie ein:

$$\begin{aligned} M &:= \begin{pmatrix} -1 \cdot -1 \cdot \langle (1,1),(1,1) \rangle & -1 \cdot 1 \cdot \langle (1,1),(2,2) \rangle \\ 1 \cdot -1 \cdot \langle (2,2),(1,1) \rangle & 1 \cdot 1 \cdot \langle (2,2),(2,2) \rangle \end{pmatrix} \\ &= \begin{pmatrix} 2 & -4 \\ -4 & 8 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} L(\vec{\alpha}) &= 2\alpha - \frac{1}{2}\left(\begin{pmatrix} \alpha & \alpha \end{pmatrix}^T \begin{pmatrix} -2\alpha & 4\alpha \end{pmatrix}\right) \\ &= 2\alpha - \frac{1}{2}\left(-2\alpha^2 + 4\alpha^2\right) \\ &= -\alpha^2 + 2\alpha \end{aligned}$$

1. Ableitung: $L'(\vec{\alpha}) = -2\alpha + 2$, Nullstelle $\alpha = 1$.
Ist das gesuchte Maximum, da es sich hier offenbar um eine umgedrehte Parabel handelt.
Daher Lösung: $\alpha_1 = \alpha_2 = 1$

(b) Bestimmung des Normalenvektors:

$$\begin{aligned}\vec{w} &= \sum_{i=1}^{n} \alpha_i \cdot y_i \cdot \vec{x}_i \\ &= 1 \cdot -1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \cdot 1 \cdot \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}\end{aligned}$$

(c) Bestimmung des Skalars b: Anmerkung: $\min$, $\max$ hier trivial, da je nur ein Support Vector!

$$\begin{aligned}b &= -\frac{\max_{i,y_i=-1} \langle \vec{w}, \vec{x}_i \rangle + \min_{i,y_i=1} \langle \vec{w}, \vec{x}_i \rangle}{2} \\ &= -\frac{\begin{pmatrix} 1 \\ 1 \end{pmatrix}^T \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}^T \cdot \begin{pmatrix} 2 \\ 2 \end{pmatrix}}{2} \\ &= -\frac{2+4}{2} \\ &= -3\end{aligned}$$

(d) Klassifikation:

Entscheidungsregel:

$$h(\vec{x}) = sig\left(\sum_{x_i \in SV} \alpha_i \cdot y_i \langle \vec{x_i}, \vec{x} \rangle + b\right)$$

$$\begin{aligned}h((3,5)) &= sig\left(1 \cdot -1 \langle (1,1),(3,5) \rangle + 1 \cdot 1 \langle (2,2),(3,5) \rangle - 3\right) \\ &= sig(-8 + 16 - 3) \\ &= sig(5) \\ &= +1 \\ &= y_2\end{aligned}$$

$$\begin{aligned}h((0,1)) &= sig\left(1 \cdot -1 \langle (1,1),(0,1) \rangle + 1 \cdot 1 \langle (2,2),(0,1) \rangle - 3\right) \\ &= sig(-1 + 2 - 3) \\ &= sig(-2) \\ &= -1 \\ &= y_1\end{aligned}$$