Übung zur Vorlesung

Maschinelles Lernen 2

Sommersemester 2011

Abteilung Maschinelles Lernen
Institut für Softwaretechnik und
theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Blatt 8

Abgabe bis Montag, 20. Juni 2011, 12:00 Uhr, Briefkasten bei Raum FR6061

In diesem Übungsblatt geht es um die Themen "Konvexe Optimierung" und "Multiple-Kernel Learning". Ingesamt können 40 Punkte erzielt werden.

1. **Geometrische Interpretation (10 Punkte)** Betrachte das folgende Optimierungsproblem:

$$\max_{\boldsymbol{w}\in\mathbb{R}^n,\lambda\in\mathbb{R}} \lambda \qquad \text{so dass} \qquad \begin{array}{rcl} \|\boldsymbol{w}\|^2 & \leq & 1\,, \\ \lambda & \leq & \boldsymbol{w}^T\boldsymbol{x}_i\,, \quad i = 1,\ldots,m\,. \end{array} \tag{1}$$

wobei $\boldsymbol{x}_i \in \mathbb{R}^n$, $i = 1,\ldots,m$, Vektoren sind, so dass es ein $\hat{\boldsymbol{w}}$ gibt, welches $\hat{\boldsymbol{w}}^T\boldsymbol{x}_i > 0$ für $i = 1,\ldots,m$ erfüllt.

- Zeige, dass das Optimierungsproblem (1) konvex ist.
- Das Problem lässt sich so interpretieren, dass man zu gegebenen Punkten $\boldsymbol{x}_i$ eine Hyperebene durch den Nullpunkt sucht, die durch $\boldsymbol{w}^T\boldsymbol{x} = 0$ definiert ist. Welche Hyperebene wird durch das Optimierungsproblem beschrieben?

(*Hinweis:* Man könnte für ein zunächst festes $\boldsymbol{w}$ überlegen, wie das maximale λ aussieht. Beachte auch, dass für $\|\boldsymbol{w}\| = 1$, $\boldsymbol{w}^T\boldsymbol{x}$ den Abstand von $\boldsymbol{x}$ von der Hyperebene angibt.)

2. **Lagrange-Dualität (10 Punkte)** Wir betrachten jetzt das folgende Optimierungsproblem:

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^m} \left\| \sum_{i=1}^m \alpha_i \boldsymbol{x}_i \right\| \qquad \text{so dass} \qquad \begin{array}{rcl} \sum_{i=1}^m \alpha_i & = & 1\,, \\ \alpha_i & \geq & 0\,, \quad i = 1,\ldots,m\,, \end{array} \tag{2}$$

- Zeige, dass das Problem (2) zu Problem (1) dual ist, und zeige, dass die Dualitätslücke ("duality gap") null ist.
- Probelm (2) kann so interpretiert werden, dass man einen Punkt aus der konvexen Hülle der Punkte $\boldsymbol{x}_i$ sucht. Welcher Punkt wird durch das Problem beschrieben?

(*Hinweis:* Der Lagrangefaktor für die Nebenbedingung $\|\boldsymbol{w}\|^2 \leq 1$ verschwindet nicht sofort in der Rechnung, kann aber ganz zum Schluss explizit maximiert werden.)

3. **Lernen mit mehreren Kernen (20 Punkte).** Um mit mehreren Kernen zu Lernen kann man das folgende Optimierungsproblem lösen

$$\begin{aligned} \min_{\boldsymbol{w},\xi,\boldsymbol{\beta}} \quad & \frac{1}{2}\|\boldsymbol{w}\|_2^2 + C\sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & y_i\langle \boldsymbol{w}, \phi_{\boldsymbol{\beta}}(x_i)\rangle \geq 1 - \xi_i, \;\; \xi_i \geq 0 \\ & \boldsymbol{\beta} \geq \boldsymbol{0}, \;\; ||\boldsymbol{\beta}||_p \leq 1\,, \end{aligned} \tag{3}$$

wobei $\phi_{\boldsymbol{\beta}} = \sqrt{\beta_1}\phi_1 \times \cdots \times \sqrt{\beta_m}\phi_m$ und der Gewichtsvektor $\boldsymbol{w}$ hat die Blockstruktur $\boldsymbol{w} = (\boldsymbol{w}_1^\top,\ldots,\boldsymbol{w}_m^\top)^\top$.

(a) Das Optimierungsproblem (3) ist nicht konvex. Finde eine konvexe Formulierung des Problems und argumentiere, dass diese tatsächlich konvex ist. (*Hinweis:* Ersetze den Gewichtsvektor geschickt durch $\boldsymbol{w}_j^{\text{NEW}} = \sqrt{\beta_j}\boldsymbol{w}_j^{\text{OLD}}$. Siehe auch Kapitel 3 in `http://www.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf`

(b) Betrachte das aus (a) resultierende Optimierungsproblem. Nimm weiterhin an, dass die Variablen $\boldsymbol{w}$ und $\boldsymbol{\xi}$ fest sind, d.h. $\boldsymbol{\beta}$ die einzige freie Variable ist. Zeige, dass das Optimum des Problems bei

$$\beta_j^* = \frac{\|\boldsymbol{w}_j\|^{\frac{2}{p+1}}}{\sqrt[p]{\sum_{l=1}^{m} \|\boldsymbol{w}_l\|^{\frac{2p}{p+1}}}}$$

erreicht wird. (*Hinweis:* Nutze die KKT-Bedingungen für Deinen Beweis.)

Für Fragen zum Übungsblatt bitte in der Google Group http://groups.google.com/group/ml-tu registrieren und die Fragen an die Mailingliste stellen.