

Maschinelles Lernen 2

Sommersemester 2008

Blatt 8

Abgabe 10. Juni 2008 in der Vorlesung. Praktische Übungsaufgaben über PASS abgeben (<https://ml01.zrz.tu-berlin.de/~mikio/pass.pl?conf=blatt8.conf>). Verwende matlab (/home/ml/ml/bin), oder octave (<http://www.octave.org> frei verfügbar).

Aufgaben

In der Vorlesung wurden Large-Scale-Lernmethoden anhand von Support Vector Machines besprochen. In der Übung soll die der stochastische Gradienten Abstieg weiter vertieft werden, da er besonders einfach zu implementieren und in der Praxis erstaunlich gute Resultate liefert.

1. Zeige, dass das primale SVM Problem mit Nebenbedingungen

$$\min_{\mathbf{w} \in \mathbf{R}^D, b \in \mathbf{R}, \xi \in \mathbf{R}^N} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i$$

so dass $\xi_i \geq 0, \quad 1 \leq i \leq n$
 $y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \quad 1 \leq i \leq n$

äquivalent zu

$$\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N (\max\{0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)\})$$

ist. (10 Punkte)

2. Leite die Update-Regeln für Stochastic Gradient Descent für quadratische Verlustfunktionen

$$\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N (\max\{0, 1 - y_i(\mathbf{w}^T \mathbf{x}_i)\})^2$$

her. (10 Punkte)

3. Implementiere Stochastic Gradient Descent für die Verlustfunktion in Aufgabe 2 in einer Funktion `sgd.m` die die Argumente `X`, `Y`, und `C`, und Anzahl Epochen und Anzahl der Iterationen pro Epoche übergeben wird (Zur Erinnerung: In jeder Epoch wird die Lernrate wieder auf den Initialwert gesetzt, pro Epoche läuft die Optimierung dann eine gewisse Anzahl von Iterationen).

Schreibe ein Skript `alpha_dataset.m`, welches den `alpha` Datensatz (zu finden auf der Vorlesungsseite) lädt, trainiert, und die Labels auf dem Testdatensatz vorhersagt. Wähle Parameter durch Ausprobieren auf einem Teil der Daten, eine vollständige Kreuzvalidierung ist nicht erforderlich. Gib die prädizierten Label ebenfalls als `predicted_labels.txt` ab. (20 Punkte)