MachineLearning for Visual Computing
Aufgabenblock 1

Christian Brändle

November 23, 2014

Abstract

0.1 Einfaches Perceptron - Datengeneration

Gegeben sind vier Datensets a 100 Beobachtungen von 2-dimensionalen Eingangsdaten, welche normalverteilt sind. Die Figuren in Tabelle 1 verdeutlichen dies.

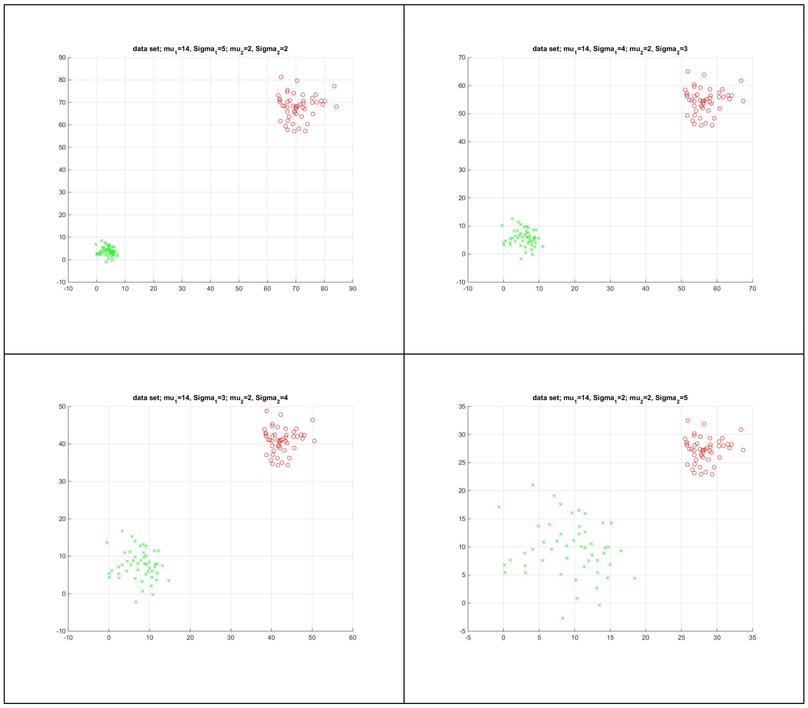


Table 1: Iteration distribution over gamma

0.2 Einfaches Perceptron - Perceptrontraining

0.2.1 Anzahl Iterationen

Die Anzahl der Iterationen hängt maßgeblich von der *margin* der separierbaren Daten ab. Je kleiner die margin, desto mehr Iterationen sind notwendig, bis der Algorithmus terminiert.

0.2.2 Welchen Einfluß hat die Schrittweite

Der Einfluß der Schrittweite hängt sowohl von den Eingangsdaten als auch vom gewählten Algorithmus ab. Es ist zu beobachten, daß der batch-Algorithmus keine großen Abweichungen zeigt, egal welches γ gewählt wird. Beim online-Verfahren ist der Einfluß größer, das heißt die Varianz in den ermittelten Iterationen ist höher, aber auch hier ist kein eindeutiger Bereich über alle Simulationen auszumachen, wo ein gegebenes γ die Iterationen des Algorithmus wesentlich reduziert. Zu sehen ist die Verteilung der Iterationen über ein γ von 0.1 bis 4 für die jeweiligen unterschiedlichen Datensätze in Tabelle 2.

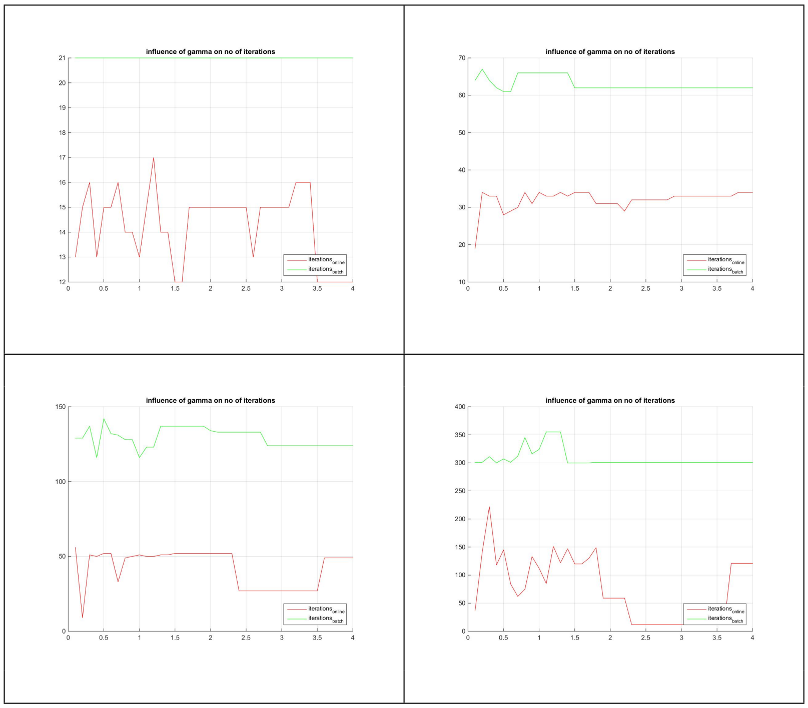


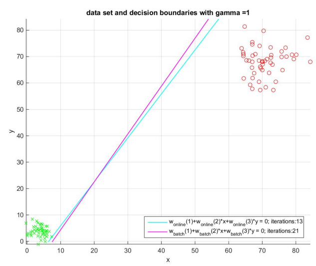
Table 2: Iteration distribution over gamma

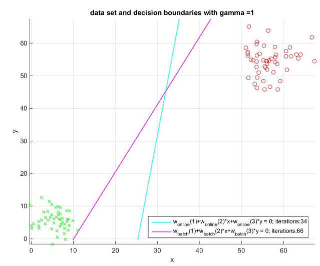
0.2.3 Daten und Entscheidungsgrenzen im $\mathbb{R}^2$

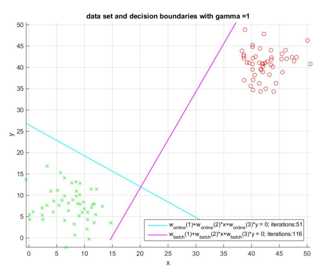
Die ermittelten Entscheidungsgrenzen des batch-learning sowie des online-learning Algorithmus sind für die einzelnen Datensätze und ein gegebenes γ von 1 in Tabelle 3 zu sehen.

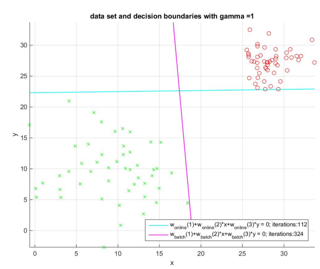
0.2.4 Wie ist das Verhalten bei nicht linear separierbaren Daten

Der Algorithmus terminiert nicht da bei jedem Durchlauf Daten gefunden werden, welche in der falschen Klasse landen.

<table>
<tr>
<td>

set_1: $\mu_1 = 10$, $\Sigma_1 = 5$
set_2: $\mu_2 = 2$, $\Sigma_2 = 2$</td>
<td>

set_1: $\mu_1 = 10$, $\Sigma_1 = 4$
set_2: $\mu_2 = 2$, $\Sigma_2 = 3$</td>
</tr>
<tr>
<td>

set_1: $\mu_1 = 10$, $\Sigma_1 = 3$
set_2: $\mu_2 = 2$, $\Sigma_2 = 4$</td>
<td>

set_1: $\mu_1 = 10$, $\Sigma_1 = 2$
set_2: $\mu_2 = 2$, $\Sigma_2 = 5$</td>
</tr>
</table>

Table 3: Iteration distribution over gamma

0.3 Lineare Regression

0.3.1 Gewichtsvektor mittels Gradientenabstieg bei quadratischer Fehlerfunktion

0.3.2 Optimaler Gewichtsvektor w^*

Den optimalen Gewichtsvektor w^* kann man über die Bestimmung der Pseudoinversen A^+ für die Gleichung $Aw^* = b$ ermitteln. Die Gleichung $w^* = A^+b$ bestimmt das optimale w^*. Dabei berechnet sich A^+ wie folgt:

$$A^+ = (AA^T)^{-1} \quad \text{für invertierbare } AA^T \tag{1}$$

$$A^+ = (AA^T)^{-1} + \lambda I \quad \text{für nicht invertierbare} AA^T \text{ mit } \lambda << 1 \tag{2}$$

Das ermittelte Ergebnis für w^* mit normalisierten Daten lautet: $w^* = (1.0106678, -10.0086, 1.8960336, 0.0262634)$.

0.3.3 Konvergenzverhalten in Abhängigkeit von γ