Übung Neuroinformatik - Blatt 09

Gruppe AC

5. Juli 2019

Aufgabe 1: Gewichtsinitialisierung

Teilaufgabe 1

a)

$$
\begin{aligned}
F(u_i) &= \sum_{k=1}^{m} x_k \cdot \mathcal{N}\{w_{ki}^{(1)}\} \quad , m = 300 \\
&= \sum_{k=1}^{m/2} 1 \cdot \mathcal{N}(0, 1) \\
&= \mathcal{N}(0, \sqrt{150})
\end{aligned}
$$

b) Die Aufsummierung der Standardnormalverteilungen führt zu der in Abbildung 1 dargestellten Normalverteilung. Diese ist beinahe gleichverteilt, was mit zunehmendem m verstärkt wird. Da die Gewichte alle sehr unterschiedlich sind, fällt der zugehörige Gradient ebenfalls sehr unterschiedlich aus. Unter Umständen konvergieren manche Gewichte deutlich langsamer als andere.

c)

$$
\begin{aligned}
F(u_i) &= \sum_{k=1}^{m} x_k \cdot \mathcal{N}\{w_{ki}^{(1)}\} \quad , m = 300 \\
&= \sum_{k=1}^{m/2} 1 \cdot \mathcal{N}(0, 1/\sqrt{m} \\
&= \mathcal{N}\left(0, \sum_{k=1}^{m/2} \left(\frac{1}{\sqrt{m}}\right)^2\right) = \mathcal{N}\left(0, \sum_{k=1}^{m/2} \frac{1}{m}\right) \\
&= \mathcal{N}\left(0, 0.5\right)
\end{aligned}
$$

d) Die Aufsummierung der *truncated normal distribution* führt nicht zu einem so großen Sigma-Wert der resultierenden Verteilung. (vgl. Abbildung **??**

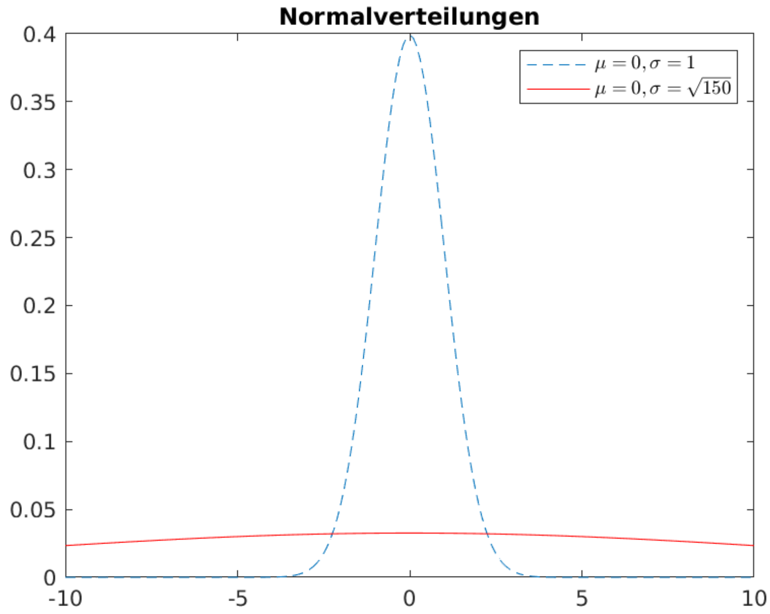

Abbildung 1: Normalverteilungen zu Teilaufgabe 1b

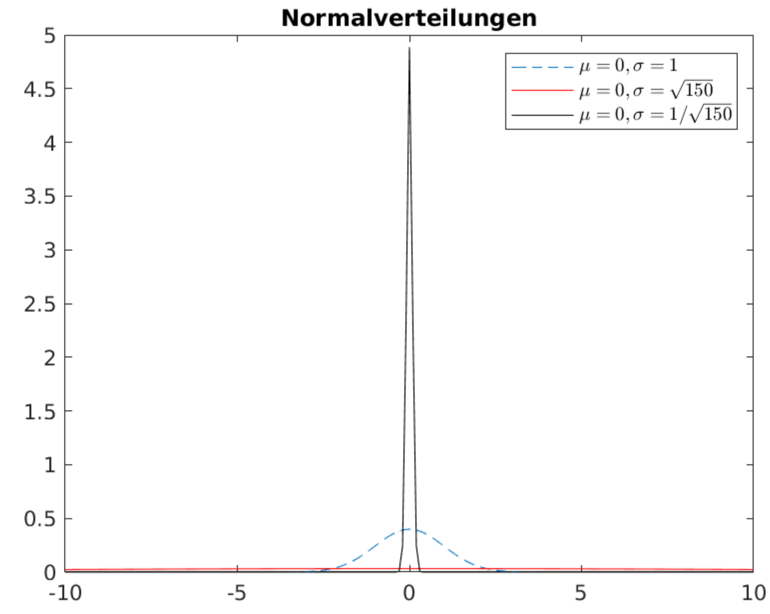

Abbildung 2: Normalverteilungen zu Teilaufgabe 1d

Teilaufgabe 2

A entspricht der skalierten Normalverteilung: Die Gewichte sind entsprechend der in Teilaufgabe 1 berechnete Normalverteilung verteilt.

B entspricht der Standardnormalverteilung: Die Gewichte und axonalen Potentiale sind nahezu gleichverteilt. (Vgl. Teilaufgabe 1b)

Teilaufgabe 3

Problem In der gegebenen Formel zur Berechnung des Gradienten kommt der Term $\sum_{i=1}^{n} w_{ij}^{(2)} \cdot x_k$ vor. Da $w_{ij}^{(2)}$ wie in Teilaufgabe 1 auch Normalverteilt ist, tritt hier der gleiche Effekt/ das gleiche Problem auf. Durch Aufsummierung addieren sich die Standardabweichungen zu einem großen Wert auf. Die Gradienten sind dann sehr unterschiedlich groß.

Lösung Durch die Xavier-Initialisierung wird analog zur Teilaufgabe 1c) die Verteilung dermaßen geschmälert, dass die Summe der Verteilungen wiederum akzeptable Kennwerte aufweist.