Deep Learning in NLP // Ha. Nr. 3 (theoretical)

Gerrit Leuteritz (2645434) Simon Münker (2611976) Helen Hense(2423533)

November 1, 2019

0.1 Gegeben:

$D \in \mathbb{R}^3 : \{(3, 0.2, 4), (5, 3.5, 3.2), (10, 4.6, 2), (12.5, 0, 1.4)\}$

0.2 Gesucht:

Ermitteln Sie durch eine lineare Regression die lineare Funktion h
$h : \mathbb{R}^2 \to \mathbb{R}$, für die gilt: für $(x, y, z) \in D$ ist die summierte Abweichung $(h(x, y) - z)^2$ minimal

0.3 Lösungsversuch:

$$argmin_{a_1,a_2,b\in\mathbb{R}} \sum_{(x,y,z)\in D} (a_1x + a_2y + b - z)^2 \tag{1}$$

Da unser Datensatz endlich viele Punkte hat, versuchen wir folgende Funktion zu minimieren:

$$\begin{aligned}
&\sum_{(a_1,a_2,b)\in D} (a_1x + a_2y + z - b)^2 = && \quad (2)\\
&(3x + 0.2y + z - 4)^2 + && (D_1) \quad (3)\\
&(5x + 3.5y + z - 3.2)^2 + && (D_2) \quad (4)\\
&(10x + 4.6y + z - 2)^2 + && (D_3) \quad (5)\\
&(12.5x + 0y + z - 1.4)^2 = && (D_4) \quad (6)
\end{aligned}$$

Nach arithmetische Umformungen (binomische Formel) erhalten wir diese Funktion:

$$f(x, y, z) = 290.25x^2 + 128.2xy + 61.0xz - 131.0x + 33.45y^2 + 16.6yz - 42.4y + 4.0z^2 - 21.2z + 32.2 \tag{7}$$

Wir leiten $f(x, y, z)$ partiell ab:

$$\begin{aligned}
&580.5x + 128.2y + 61.0z - 131.0 && (f' : x) \quad (8)\\
&128.2x + 66.9y + 16.6z - 42.4 && (f' : y) \quad (9)\\
&61.0x + 16.6y + 8z - 21.2 && (f' : z) \quad (10)
\end{aligned}$$

Und bilden den dazugehörigen Gradienten:

$$\begin{aligned}
\nabla f(x, y, z) = ((&580.5x + 128.2y + 61.0z - 131.0), && (11)\\
(&128.2x + 66.9y + 16.6z - 42.4), && (12)\\
(&61.0x + 16.6y + 8z - 21.2)) && (13)
\end{aligned}$$

Wir haben also 3 Gleichungen, die wir auf 0 setzen müssen:

$$580.5x + 128.2y + 61.0z - 131.0 = 0 \tag{14}$$
$$128.2x + 66.9y + 16.6z - 42.4 = 0 \tag{15}$$
$$61.0x + 16.6y + 8z - 21.2 = 0 \tag{16}$$

Wenn wir die Gleichungen nach x, y, z auflösen erhalten wir:

$$x = -0.265152444442663 \tag{17}$$
$$y = -0.0357149061094030 \tag{18}$$
$$z = 4.74589581905231 \tag{19}$$

Daraus folgt für h:

$$h(x, y) = -0.265152444442663x + -0.0357149061094030y + 4.74589581905231 \tag{20}$$

Nun berechnen wir Unterschied zwischen unseren Daten D und den Ergebnissen von h:

$$\sum_{(x,y,z)\in D} (h(x, y) - z)^2 = 0.0181454385592173 \tag{21}$$