Prof. Dr. Morik
Christian Bockermann

20. Oktober 2008

Solution to Exercises

Machine Learning

Wintersemester 2008/2009
Exercise Sheet 2

Exercise 2 Das lineare Modell hat die Form

$$y = m \cdot x + b.$$

Die quadratische Fehlersumme RSS für gegebenes m und b ist

$$\begin{aligned}
RSS(m,b) &= \sum_{i=1}^{n} [y_i - (mx_i + b)]^2 \\
&= \sum_{i=1}^{n} \left(y_i^2 - 2y_i(mx_i + b) + (mx_i + b)^2\right) \\
&= \sum_{i=1}^{n} y_i^2 - 2\sum_{i=1}^{n} y_i(mx_i + b) + \sum_{i=1}^{n}(mx_i + b)^2 \\
&= \sum_{i=1}^{n} y_i^2 - 2\sum_{i=1}^{n}(y_i m x_i + y_i b) + \sum_{i=1}^{n} \left[m^2 x_i^2 + 2mx_i b + b^2\right]
\end{aligned}$$

Die notwendigen Bedingungen für die Existenz eines lokalen Extremums (wir suchen ein lokales/globales Minimum) sind

$$\frac{\partial RSS(m,b)}{\partial b} = 0 \text{ und } \frac{\partial RSS(m,b)}{\partial m} = 0$$

Da es sich bei RSS um eine positive, quadratische Funktion handelt, existiert nur ein lokales Minimum, welches zugleich auch globales Minimum ist.

Die beiden partiellen Ableitungen nach b und m sind

$$\frac{\partial RSS(m,b)}{\partial b} = -2\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} 2mx_i + \sum_{i=1}^{n} 2b = -2\sum_{i=1}^{n} y_i + 2\sum_{i=1}^{n} mx_i + 2nb$$

und

$$\frac{\partial RSS(m,b)}{\partial m} = -2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} \left[2mx_i^2 + 2x_i b\right] = -2\sum_{i=1}^{n} x_i y_i + 2m\sum_{i=1}^{n} x_i^2 + 2b\sum_{i=1}^{n} x_i$$

Wir setzen beide Ableitungen gleich 0:

$$\begin{aligned}
\frac{\partial RSS(m,b)}{\partial b} &= 0 \\
&\Rightarrow \quad 2nb = 2\sum_{i=1}^{n} y_i - 2\sum_{i=1}^{n} mx_i \\
&\Rightarrow \quad b = \frac{1}{n}\sum_{i=1}^{n}(y_i - mx_i) \\
&\Rightarrow \quad b = \underbrace{\frac{1}{n}\sum_{i=1}^{n} y_i}_{=\overline{y}} - m\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)}_{=\overline{x}} = \overline{y} - m\overline{x}
\end{aligned}$$

Dabei entspricht $\overline{x}$ bzw. $\overline{y}$ jeweils dem arithmetischen Mittel von x bzw. y.

$$\begin{aligned}
\frac{\partial RSS(m,b)}{\partial m} &= 0 \\
&\Rightarrow \quad 2m\sum_{i=1}^{n} x_i^2 = 2\sum_{i=1}^{n} x_i y_i - 2b\sum_{i=1}^{n} x_i \\
&\Rightarrow \quad m\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} x_i \\
&\Rightarrow \quad m\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i - (\overline{y} - m\overline{x})\sum_{i=1}^{n} x_i \\
&\Rightarrow \quad m\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i - \overline{y}\sum_{i=1}^{n} x_i + m\overline{x}\sum_{i=1}^{n} x_i \\
&\Rightarrow \quad m\sum_{i=1}^{n} x_i^2 - m\overline{x}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i - \overline{y}\sum_{i=1}^{n} x_i \\
&\Rightarrow \quad m = \frac{\sum_{i=1}^{n} x_i y_i - \overline{y}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \overline{x}\sum_{j=1}^{n} x_i}
\end{aligned}$$

Für die gegeben Datenpunkte $(x_1, y_1), \ldots, (x_9, y_9)$ ergibt sich

$$\begin{aligned} m &= \frac{258 - \frac{43}{9}45}{285 - 5 \cdot 45} \approx 0.7167 \\ b &= \frac{43}{9} - 0.7167 \cdot 5 \approx 1.194 \end{aligned}$$

Dies führt zur Regressionsgeraden

$$f(x) = 0.7167x + 1.194$$