



---

## Mathematisch-algorithmische Grundlagen für Data Science

### Lösungen zu Aufgabenblatt 6

---

#### Aufgabe 3 (Logistische Regression)

1+2+3+2+2=10 Punkte

Diese Aufgabe wiederholt im Wesentlichen die logistische Regression so, wie Sie sie in WuS kennengelernt haben, also mit einer Variablen und zwei Klassen. Auf späteren Aufgabenblättern werden wir diesen Ansatz auf mehrere Variablen und mehr als zwei Klassen verallgemeinern.

Gegeben ist eine Menge von Objekten, die jeweils zwei Merkmale haben:  $X \in \mathbb{R}$  und  $Y \in \{0, 1\}$  (die Klasse des Objekts). Das Merkmal  $Y$  können wir als Bernoulli-verteilte Zufallsvariable auffassen, deren Erfolgswahrscheinlichkeit von  $x$  abhängt.

Gesucht ist eine sigmoide Funktion

$$p(x) = \frac{1}{1 + e^{-\lambda_1 x + \lambda_2}},$$

die die Wahrscheinlichkeit  $P(Y = 1|X = x)$  schätzt, also die Wahrscheinlichkeit dafür, dass ein Objekt zur Klasse 1 gehört, unter der Bedingung, dass der Merkmalswert von  $X$  den Wert  $x$  hat.

(c) Zeigen Sie, dass die Funktion  $\text{sig} : \mathbb{R} \rightarrow (0, 1)$  mit

$$\text{sig}(x) = \frac{1}{1 + e^{-x}}$$

als Umkehrfunktion die Funktion  $\text{logit} : (0, 1) \rightarrow \mathbb{R}$  mit

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

hat.

**Lösung:** Es sei  $y = \text{sig}(x)$ , also

$$y = \frac{1}{1 + e^{-x}}.$$

Um die Umkehrfunktion zu berechnen, vertauschen wir  $x$  und  $y$ :

$$x = \frac{1}{1 + e^{-y}}.$$

Diese Gleichung lösen wir nach  $y$  auf:

$$\begin{aligned}x &= \frac{1}{1 + e^{-y}} && \Leftrightarrow x(1 + e^{-y}) = 1 \\&&& \Leftrightarrow 1 + e^{-y} = \frac{1}{x} \\&&& \Leftrightarrow e^{-y} = \frac{1}{x} - 1 \\&&& \Leftrightarrow -y = \log\left(\frac{1-x}{x}\right) \\&&& \Leftrightarrow y = -\log\left(\frac{1-x}{x}\right) \\&&& \Leftrightarrow y = \log\left(\left(\frac{1-x}{x}\right)^{-1}\right) \\&&& \Leftrightarrow y = \log\left(\frac{x}{1-x}\right)\end{aligned}$$