

Errata

(Mathematische Einführung in Data Science von Sven-Ake Wegner)

16. März 2024

- Seite 17, Zeile 11: r_{xy}
- Seite 24, Zeile -11: $f = \langle (a_1, \dots, a_d), \cdot \rangle + a_0$
- Seite 26, Zeile 18: $f(\cancel{z}) = \text{sig}(\langle w, \cdot \rangle)$
- Seite 31, Zeile 12: \dots und $\langle w, \widehat{w}_k \rangle < 0$ gelten.
- Seite 38, Zeile 13: \dots und $x_1, \dots, x_k \in D_1 := \{x \mid (x, y) \in D\}$ gelten.
- Seite 38, Zeile 14: $x_1 \in \underset{z \in D_1}{\text{argmin}} \rho(x, z)$ sowie $x_j \in \underset{z \in D_1 \setminus \{x_1, \dots, x_{j-1}\}}{\text{argmin}} \rho(x, z)$ für $j \geq 2$
- Seite 38, Zeile 20: $f(x) \in \underset{y \in Y}{\text{argmax}} N(y)$
- Seite 39, Zeile 4: $z^* \leftarrow \underset{z \in D'_1}{\text{argmin}} \rho(x, z)$
- Seite 39, Zeile 5: $D'_1 \leftarrow D'_1 \setminus \{(z^*, y^*)\}$
- Seite 39, Zeile 10–11: Hierbei bezeichnet $\pi_2(x, y) = x$ die Projektion auf den zweiten Eintrag von $(x, y) \in D$ und y^* das Label von z^* .
- Seite 39, Zeile -2: \dots vorgenannten grauen Punkt x mit \dots
- Seite 41, Zeile 6–7: \dots Trainingsdaten D_{Train} und Testdaten D_{Test} . Dann bestimmt man einen Klassifizierer $f: X \rightarrow Y$ anhand von D_{Train} , und stellt fest, welcher Anteil der Punkte aus D_{Test} durch f korrekt klassifiziert wird.
- Seite 52, Zeile -14: Text D
- Seite 55, Zeile 21–23: \dots bevor der minimale Abstand zwischen den Clustern in der nächsten Runde erstmalig über einen einzugebenden Wert $\delta > 0$ wachsen würde.
- Seite 55, Zeile -15: D
- Zeile -12: **while** $\min_{i \neq j} \rho(C_i, C_j) \leq \delta$ **do**
- Seite 56, Zeile 4: \dots , wie in Definition 4.2, \dots
- Seite 56, Zeile 8: Aufgabe 4.2
- Seite 70, Zeile 3–4: Wir wählen $v_1 = \mathbb{1}$ als Eigenvektor zum Eigenwert λ_1 . Dann gelten für Eigenwert und Eigenvektor λ_2 bzw. v_2
- Seite 70, Zeile 7: $v_2 \in \dots$
- Seite 71, Zeile 5: $\emptyset \neq S \subset V$
- Seite 72, Zeilen 6 und 8: $\emptyset \neq S \subset V$
- Seite 74, Zeile 15: $\emptyset \neq S \subset V$
- Seite 74, Zeile 16: $\text{Cheeg}(G) = \frac{\#\partial S}{\text{vol } S}$
- Seite 102, Zeile 17: $r = \text{rk}(A)$
- Seite 110, Zeile 7: $\mathbb{R}_{\mathcal{F}}^5$
- Seite 112, Zeile 5: $(\text{Film})_{\mathcal{F}} \quad \mathbb{R}_{\mathcal{F}} \xrightarrow{A} \mathbb{R}_{\mathcal{B}} \quad (\text{Bewerterin})_{\mathcal{B}}$
- Seite 111, Zeile 4: \mathbf{R} (statt R)
- Seite 112, Zeilen 14 und 25: \mathbf{R} (statt R)
- Seite 112, Zeile -3: \dots die Daten erst einem geeigneten Pre- oder Postprocessing zu unterziehen \dots
- Seite 129, Zeile -1: $0 < \varepsilon < 1$
- Seite 134, Zeile 9: Für solche Zufallsvektoren erhält man \dots
- Seite 139, Zeile -1: $\|X^{(i)}\| \approx 1$
- Seite 146, Zeile 5: $P[\|\mathbf{X}\| - \sqrt{d} \geq \varepsilon]$
- Seite 167, Zeile 5–7: ~~Im Fall der Varianz sind diese eher technisch, und wir formulieren daher im Satz für die Varianz nur die sich ergebende qualitative asymptotische Aussage.~~

- Seite 183, Zeile −1: Wenn D linear trennbar ist, ...
- Seite 186, Zeile 8: $\text{---return } w^{(j)}$
- Seite 189, Zeile −5: $(x, y) \in D$
- Seite 190, Zeile 14–15: ... die Worte „Bonus“, „Vertrag“, „das“ und „Mensa“ ...
- Seite 194, Zeilen −2 und −1: $\mathcal{R}(D), \mathcal{K}(D)$
- Seite 194, Zeile −1: $h: \mathbb{R}^d \rightarrow \mathbb{R}^d$
- Seite 195, Zeilen 2, 4, 5, 9, 10, 12, 15 und 20: $\mathcal{R}(D), \mathcal{K}(D)$
- Seite 196, Zeilen 6–8, 12 und 13: $\mathcal{R}(D), \mathcal{K}(D)$
- Seite 198, Zeile 6: $(w^*, b^*) \in M$
- Seite 198, Zeile 11–16: Es folgt also nach Proposition 17.14, dass $w_1^* = w_2^* =: w^*$ ist. Gelte ohne Einschränkung $b_1^* < b_2^*$. Dann wählen wir für (w^*, b_1^*) und (w^*, b_2^*) jeweils i_1 und einen Index i wie in Teil ② und erhalten

$$y_i(\langle w^*, x_i \rangle + b_1^*) = \langle w^*, x_i \rangle + b_1^* < \langle w^*, x_i \rangle + b_2^* = 1$$

im Widerspruch dazu, dass (w^*, b_1^*) in M liegt. Ist $b_1^* > b_2^*$, so vertauschen wir die Rollen von i_1 und i_2 .

- Seite 198, Zeilen −11, −8 und −4: $\mathcal{R}(D), \mathcal{K}(D)$
- Seite 200, Zeilen 7: $L(x, \theta, \mu) := f(x) - \sum_{i=1}^q \theta_i g_i(x) + \sum_{j=r}^q \mu_j h_j(x)$
- Seite 200, Zeilen 18: ..., ist x^* ein Minimierer des oben angegebenen Optimierungsproblems ...
- Seite 203, Zeilen 17: ... ist es möglich, i_0 mit $\lambda_{i_0} \neq 0$...
- Seite 208, Zeilen 13: ... und Satz 14.12(ii) ist [GK02, Satz 2.46].
- Seite 209, Zeilen −6: ..., wenn man einen der Punkte x_6^*, x_7^* , oder x_{11}^* weglässt?
- Seite 216, Zeile −14: ..., sodass $y_i(\langle w, x_i \rangle + b) \geq 1$ für ...
- Seite 220, Zeile 16: ... mit $\lambda_{i_0}^* \neq 0$ ist ...
- Seite 220, Zeile 18: ... Wahl von λ^* und ...
- Seite 224, Zeile 6: Haben wir also $\langle f, f \rangle = 0$ für alle $x \in X$, so ist ...
- Seite 224, Zeile −10: ... positive Semidefinitheit der Gram-Matrix ...
- Seite 226, Zeile 12: Geben Sie an, welche $x \in \mathbb{R}$ vom zurückgezogenen Klassifizierer ...
- Seite 230, Zeile 15: die Rectified Linear Unit
- Seite 231, Zeile −4: ... Neuronen aus Proposition 16.2 zusammensetzen
- Seite 232, Zeile 3: ... sonst,
- Seite 236, Zeile 9: B ist die Matrix ohne das $(\cdot)^{-1}$.
- Seite 239, Zeile 15: $\mathcal{F}(f_1 * f_2) = \mathcal{F}f_1 \cdot \mathcal{F}f_2$
- Seite 262, Zeile 1–4: Den Rieszschen Darstellungssatz in der Version 16.19 kann man in [Wer18, Theorem II.2.5] nachlesen, die andere Version 16.16, oft auch Riesz-Markov-Theorem genannt, findet man in [Rud87, Theorem 6.19].
- Seite 270, Zeile 14: Dann gilt $h'(t) = \langle \nabla f(y + t(x - y)), x - y \rangle$...
- Seite 270, Zeile 16: $\dots = \int_0^1 \langle \nabla f(y + t(x - y)), x - y \rangle dt$
- Seite 270, Zeile 18: $\dots = \int_0^1 \langle \nabla f(y + t(x - y)), x - y \rangle - \langle f(y), x - y \rangle dt$
- Seite 271, Zeile 14: Folgerung 17.10. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar und konvex.
- Seite 275, Zeile −2: Dann gilt wegen
- Seite 283, Zeile −9: 4: $x^{(k+1)} \leftarrow x^{(k)} - \gamma_k \nabla f(x^{(k)})$.
- Seite 296, Zeile −14: S. Shalev-Shwartz und S. Ben-David, Understanding