

Zeit	Raum	Abgabe im Moodle; Mails mit Betreff: [SMD1718]
Di. 10-12	CP-03-150	philipp2.hoffmann@udo.edu und jan.soedingrekso@udo.edu
Di. 16-18	P1-02-110	felix.neubuerger@udo.edu und tobias.hoinka@udo.edu
Di. 16-18	CP-03-150	simone.mender@udo.edu und maximilian.meier@udo.edu

Aufgabe 22: *Lineare Klassifikation mit Softmax*

10 P.

Für eine Parameteranpassung bei der Klassifikation mittels der Softmax-Funktion muss der Gradient der Kostenfunktion für alle anzupassenden Parameter bestimmt werden. Die Kostenfunktion C ist wie aus der Vorlesung bekannt, gegeben durch:

$$C(f) = \frac{1}{m} \sum_{i=1}^m \hat{C}(f_i) = \frac{1}{m} \sum_{i=1}^m \left[- \sum_{k=1}^K \mathbb{1}(y_i = k) \log \frac{\exp(f_{k,i})}{\sum_j \exp(f_{j,i})} \right]. \quad (1)$$

Zur Ableitung der Kostenfunktion wird die Kettenregel verwendet:

$$\nabla_W \hat{C} = \sum_{k=1}^K \frac{\partial \hat{C}}{\partial f_{k,i}} \cdot \frac{\partial f_{k,i}}{\partial W} \quad (2)$$

$$\nabla_b \hat{C} = \sum_{k=1}^K \frac{\partial \hat{C}}{\partial f_{k,i}} \cdot \frac{\partial f_{k,i}}{\partial b}. \quad (3)$$

- a) Gegeben seien K Klassen und m Beispiele x_i jeweils mit M Komponenten. Welche Dimension haben die einzelnen Komponenten x_i , C , W , b , $\nabla_W \hat{C}$, $\nabla_{f_i} \hat{C}$, $\frac{\partial f_{k,i}}{\partial W}$, $\frac{\partial f_{k,i}}{\partial b}$? Unterscheiden Sie dabei auch zwischen Zeilen- und Spaltenvektoren.
- b) Zeigen Sie, dass sich für die Ableitung der Kostenfunktion nach den Scores für die Klasse a folgendes ergibt:

$$\nabla_{f_a} C(f) = \frac{1}{m} \sum_{i=1}^m \left[\frac{\exp(f_{a,i})}{\sum_j \exp(f_{j,i})} - \mathbb{1}(y_i = a) \right]. \quad (4)$$

- c) Bestimmen Sie als zweiten Schritt der Kettenregel die Ableitungen von $f_{k,i}$ nach W und b mit $f_{k,i} = W_k x_i + b_k$.
- d) Implementieren Sie die lineare Klassifikation mit Softmax für die zwei Populationen P0 und P1 ¹. Verfahren Sie dabei wie folgt:

- Lesen Sie die Populationen mit `PX = np.load('PX.npy')` ein.

¹Blatt04 Aufgabe1, es befinden sich numpy-Dateien im Moodle

- Führen Sie beide Populationen zusammen und erstellen Sie die entsprechenden Label (P1 hat dabei Label 1).
- Initialisieren Sie die Gewichtsmatrix und den Bias-Vektor.
- Nutzen Sie eine learning-rate von 0,5 und trainieren Sie 100 Epochen.
- Implementieren Sie die folgenden Schritte vektorisiert (Nutzen Sie `np.matmul`).
 - Implementieren Sie die Softmax Funktion und die Indikatorfunktion.
 - Iterieren Sie über die Anzahl der Epochen. Berechnen Sie in jeder Iteration die Softmax-Funktion für die aktuellen Parameter W und b . Bestimmen Sie mithilfe der erhaltenen Werte den Gradienten der Kostenfunktion bzgl. W und b . Und führen Sie schließlich ein Parameterupdate durch.

Tipp: Der Gradient nach W lässt sich auch nach $\nabla_W C = \nabla_f C \cdot x_i^T$ berechnen.

- e) Stellen Sie das Resultat (die trennende Gerade) zusammen mit den beiden Populationen in einem Scatterplot dar. Zur Herleitung der Geradengleichung nutzen Sie die Bedingung $f_1 = f_2$. Warum gilt das?

Aufgabe 23: Fehlerfortpflanzung

5 P.

Die Parameter einer Ausgleichsgeraden $y = a_0 + a_1 x$ wurden zu $a_0 = 1,0 \pm 0,2$ und $a_1 = 1,0 \pm 0,2$ bestimmt. Der Korrelationskoeffizient ist $\rho = -0,8$. Bestimmen Sie die Unsicherheit eines Wertes y als Funktion von x .

- a) Bestimmen Sie das Resultat analytisch sowohl unter Berücksichtigung der Korrelation als auch unter Vernachlässigung der Korrelation.
- b) Bestimmen Sie das Resultat numerisch mit einer Monte Carlo Simulation. Visualisieren Sie die Parameter a_0 und a_1 in einem Scatter-Plot.
- c) Bestimmen Sie die Vorhersagen y (Mittelwert und Standardabweichung) für feste $x = -3, 0, +3$ numerisch sowie analytisch und vergleichen Sie diese.

Aufgabe 24: *DeepLearning Kurzfragen*

5 P.

- a) Was beschreibt die Lossfunktion und wofür wird sie benötigt?
- b) Wie kann die Lossfunktion minimiert werden?
- c) Welche Funktion haben die Aktivierungsfunktionen bzw. welches Problem wird durch diese gelöst? Nennen Sie drei gängige Aktivierungsfunktionen.
- d) Was ist ein Neuron?
- e) Nennen Sie drei Anwendungsbeispiele für Neuronale Netze und beschreiben Sie kurz warum sie für diese Beispiele besonders geeignet sind.