Übung Künstliche Neuronale Netzwerke
Wintersemester 2015/2016

① Abgabe am 25. 11. 2015 bis Ende der Übung handschriftlich auf Papier (1x pro Person)
② Abgabe bis 02. 12. 2015, 23:55 Uhr im Moodle-Kurs (1x pro Gruppe)

2. Übungsblatt vom 18. 11. 2015

2.1 Regression

Eine Messreihe ist gegeben durch eine Folge von Eingabewerten $x_1, \ldots, x_P$ und Sollwerten $t_1, \ldots, t_P$ (Trainingsbeispiele). Aus dem Versuchsaufbau vermuten Sie, dass es sich um eine lineare Überlagerung einer Sinus- und einer Kosinusfunktion handelt:

$$y(\underline{w}, x) = w_1 \cos(x) + w_2 \sin(x)$$

a) Definieren Sie eine quadratische Fehlerfunktion $E(\underline{w})$. Geben Sie Formeln an, mit denen aus den gegebenen Trainingsbeispielen (x_p, t_p) die optimalen Gewichte w_1 und w_2 berechnet werden können. Lösen Sie dazu das durch die Minimierung der Fehlerfunktion entstehende Gleichungssystem. 5P ①

Hinweis: Um die Formeln übersichtlich zu halten, können Sie an geeigneter Stelle folgende Substitutionen verwenden:

$$\begin{aligned} a_{11} &= \sum_{p=1}^{P} \cos^2(x_p) \\ a_{22} &= \sum_{p=1}^{P} \sin^2(x_p) \\ a_{12} = a_{21} &= \sum_{p=1}^{P} \sin(x_p)\cos(x_p) \\ b_1 &= \sum_{p=1}^{P} t_p \cos(x_p) \\ b_2 &= \sum_{p=1}^{P} t_p \sin(x_p) \end{aligned}$$

b) Nach der Berechnung des minimalen Fehlers auf der Trainingsmenge stellen Sie fest, dass die Modellkomplexität zu klein war.

(i) Wie kann die Modellkomplexität mit trigonometrischen Funktionen erhöht werden? Mit welchem Lösungsansatz werden dann die optimalen Gewichte berechnet?

(ii) Kann eine höhere Modellkomplexität den Fehler auf der Trainingsmenge vergrößern (theoretisch/praktisch)? Kann er theoretisch zu 0 gemacht werden?

(iii) Was passiert mit dem Fehler auf der Testmenge, wenn die Modellkomplexität immer weiter erhöht wird (theoretisch und praktisch)?

Begründen Sie Ihre Antworten. 3P ①

bitte wenden

2.2 Normalverteilung

Gegeben sei eine zweidimensionale Zufallsvariable $X = (X_1, X_2)^{\mathrm{T}}$ mit $x_1 \in [-1, 1]$ und $x_2 \in [-3, 1]$, wobei alle Werte innerhalb dieses Rechtecks gleich wahrscheinlich sind.

a) Geben Sie die Wahrscheinlichkeitsdichten von X_1, X_2 und X an. Berechnen Sie Erwartungswerte und Standardabweichungen in X_1 und X_2, sowie den Erwartungswert von $X_1 \cdot X_2$. 4P ①

b) Schreiben Sie eine Funktion, die zufällige Vektoren $\underline{x}$ aus $[-1, 1] \times [-3, 1]$ erzeugt. Schreiben Sie eine weitere Funktion, die als neue Zufallsvariable den Durchschnitt $Y = \frac{1}{N} \sum_{n=1}^{N} \underline{x}_n$ aus einer Stichprobe mit N Werten der Zufallsvariablen X berechnet. 2P ②

c) Erzeugen Sie eine große Zahl $\underline{y}_1, \ldots, \underline{y}_S$ von Werten dieser Zufallsvariablen Y mit Stichprobengröße $N = 10$. Schreiben Sie eine Funktion, die aus diesen Werten den Mittelwert $\underline{\hat{\mu}} = (\hat{\mu}_1, \hat{\mu}_2)^{\mathrm{T}}$ und die Standardabweichungen $\hat{\sigma}_1$ und $\hat{\sigma}_2$ in x_1 bzw. x_2 der Zufallsvariablen Y schätzt. 4P ②

d) Erzeugen Sie ein zweidimensionales Histogramm mit Schrittweite $\Delta = 0.1$ und tragen Sie die Vektoren $\underline{y}_1, \ldots, \underline{y}_S$ ein. Dividieren Sie die Histogrammwerte durch $S \cdot \Delta^2$. 3P ②

Dieses Histogramm ist eine Annäherung an die Normalverteilung

$$\mathcal{N}(\underline{z}; \underline{\mu}, \sigma_1, \sigma_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\left(\frac{(z_1 - \mu_1)^2}{\sigma_1^2} + \frac{(z_2 - \mu_2)^2}{\sigma_2^2}\right)\right]$$

e) Schreiben Sie eine Funktion, die die Normalverteilung mit den Parametern $\underline{\hat{\mu}}$ sowie $\hat{\sigma}_1$ und $\hat{\sigma}_2$ berechnet. Dabei bezeichnen $\underline{\hat{\mu}}, \hat{\sigma}_1, \hat{\sigma}_2$ die Schätzungen aus Aufgabe c). Bestimmen Sie die Abweichung des Histogramms von der Normalverteilung durch Berechnung des quadratischen Fehlers E^{SSE}. Vergleichen Sie dazu den Wert jedes Histogrammbins mit dem Wert der Normalverteilung im Zentrum $\underline{z}$ des entsprechenden Histogrammbins. 5P ②