Universität Salzburg		Florian Graf
	Machine Learning	
Übungsblatt **7**		24 Punkte

Aufgabe 1. *Gradienten* 4 P.

Berechnen Sie den Gradienten und die Hessematrix der folgenden Funktionen $\mathbb{R}^d \to \mathbb{R}$. Berechnen Sie zunächst die partiellen Ableitungen und schreiben Sie anschließend den Gradienten (bzw. die Hessematrix) als Matrix-Vektor-Produkt.

(a) $f_1(x) = \mathbf{w}^\top \mathbf{x}$ mit $\mathbf{w}, \mathbf{x} \in \mathbb{R}^d$.

(b) $f_2(x) = \frac{1}{2}\|\mathbf{x}\|^2$ mit $\mathbf{x} \in \mathbb{R}^d$.

(c) $f_3(x) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}$ mit $\mathbf{x} \in \mathbb{R}^d$ und $A \in \mathbb{R}^{d\times d}$.

(d) $f_4(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$ mit $\mathbf{x}, \mathbf{b} \in \mathbb{R}^d$ und $A \in \mathbb{R}^{d\times d}$.

Aufgabe 2. *Lineare Regression – MLE* 10 P.

Es seien Daten $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^d$ und $(y_1, \ldots, y_n) \in \mathbb{R}$ gegeben. Wir modellieren die Zielvariable y durch ein lineares Regressionsmodell der Form $p(y_i|\mathbf{x}_i, \mathbf{w}, \sigma^2) = \mathcal{N}(y_i|\mathbf{w}^\top\mathbf{x}_i, \sigma^2)$ mit $\mathbf{w} \in \mathbb{R}^d$ und $\sigma^2 > 0$. Bestimmen Sie den gemeinsamen Maximum-Likelihood Schätzer $(\hat{\mathbf{w}}, \hat{\sigma}^2)$ der Parameter $(\mathbf{w}, \sigma^2)$.

Aufgabe 3. *Gewichtete lineare Regression* 10 P.

Es seien Daten $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^d$ und $(y_1, \ldots, y_n) \in \mathbb{R}$ gegeben. Wir betrachten ein gewichtetes lineares Regressionsmodell der Form $p(y_i|\mathbf{x}_i, \mathbf{w}) = \mathcal{N}(y_i|\mathbf{w}^\top\mathbf{x}_i, \sigma^2(\mathbf{x}_i))$ mit $\mathbf{w} \in \mathbb{R}^d$ und $\sigma^2(\mathbf{x}_i) > 0$, d.h. die Varianz der Zielgröße y hängt von der Einflussgröße $\mathbf{x}$ ab.

(a) Beschreiben Sie ein Experiment bei dem es sinnvoll ist die Beobachtungen durch ein solches gewichtetes Regressionsmodell (anstatt eines klassischen linearen Regressionsmodells) zu modellieren.

(b) Schreiben Sie die NLL des Modells mithilfe des Vektors $\mathbf{w}$, sowie den Matrizen $\mathbf{X}$ und $\mathbf{\Sigma}$, die die Einflussgrößen $\mathbf{x}_i$ bzw. die Varianzen $\sigma^2(\mathbf{x}_i)$ enthalten.

(c) Bestimmen Sie den Maximum-Likelihood Schätzer $\hat{\mathbf{w}}$ für $\mathbf{w}$.

(d) Die nachfolgende Tabelle enthält die empirischen Mittelwerte und Varianzen der Zielgröße $y \in \mathbb{R}$ für verschiedene Werte der Einflussgröße $x \in \mathbb{R}$.

x_i	y_i	$\sigma^2(x_i)$
0	0	1
1	2	1
2	2	3
4	0	3

Fitten Sie ein gewichtetes und eine ungewichtetes lineares Regressionsmodell an die Daten mittels der Maximum-Likelihood Methode. Zeichnen Sie die Datenpunkte und die resultierenden Regressionsgeraden handschriftlich in ein Diagramm ein. Vergleichen Sie die beiden Modelle.