Universität Salzburg	Florian Graf
Machine Learning	
Übungsblatt **8**	20 Punkte

Aufgabe 1. *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 $\sigma^2 > 0$.

(a) Wir nehmen an, dass σ^2 bekannt ist. Bestimmen Sie den Maximum-Likelihood Schätzer des Parameters $\mathbf{w}$. Nutzen Sie dazu die Matrix $\mathbf{X}$, deren Reihen die Vektoren $\mathbf{x}_i$ sind. Sie dürfen außerdem annehmen, dass die Matrix $\mathbf{X}^\top \mathbf{X}$ invertierbar ist.

(b) Wir nehmen an, dass $\mathbf{w}$ bekannt ist. Bestimmen Sie den Maximum-Likelihood Schätzer von σ^2.

(c) Nun sind sowohl $\mathbf{w}$ als auch σ^2 unbekannt. Bestimmen Sie den gemeinsamen Maximum-Likelihood Schätzer.

Aufgabe 2. *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 $\mathbf{w}$. Nehmen Sie dazu an, dass die $\sigma^2(\mathbf{x}_i)$ bekannt sind.

(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
3	0	3

Fitten Sie ein gewichtetes und eine ungewichtetes lineares Regressionsmodell an die Daten mittels der Maximum-Likelihood Methode (Vergessen Sie nicht den Offset bzw. Bias!). Zeichnen Sie die Datenpunkte und die resultierenden Regressionsgeraden handschriftlich in ein Koordinatensystem ein. Vergleichen Sie die beiden Modelle.