

Prof. Dr. Vladimir Shikhman
Professur für Wirtschaftsmathematik
Technische Universität Chemnitz

Übungsleiter: David Müller
david.mueller@mathematik.tu-chemnitz.de

Mathematische Grundlagen von Big Data Analytics (SS 2018)
Übung 12: Sparse Recovery II / Ridge Regression

- 1) Gegeben sei die Datenmatrix $X \in \mathbb{R}^{n,d}$, mit $n \geq d$.
- a) Geben Sie die Least Squares Lösung für $Xw = b$ an, wobei $w \in \mathbb{R}^d$ und $b \in \mathbb{R}^n$.
- b) Berechnen Sie, falls möglich, den optimalen Feature Vektor \hat{w} , wenn Ihnen folgende Daten zur Verfügung stehen:

$$X = \begin{pmatrix} -1 & 1 \\ 2 & 1 \\ 1 & -2 \end{pmatrix}, \quad b = \begin{pmatrix} 10 \\ 5 \\ 20 \end{pmatrix}$$

- c) Berechnen Sie für b), falls möglich, den optimalen Feature Vektor \hat{w} , wenn die Datenmatrix nun wie folgt aussieht:

$$\tilde{X} = \begin{pmatrix} -1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 0.5 & 0.5 \end{pmatrix}$$

- 2) Betrachten Sie das Optimierungsproblem:

$$\min_w \quad 1/2 \|Xw - b\|_2^2 + (1/2)\alpha \|w\|_2^2.$$

- a) Leiten Sie die optimale Lösung des Problems her.
- b) Wann existiert eine eindeutige Lösung?
- c) Welche Rolle spielt der Parameter α für die Lösung?
- 3) Berechnen Sie die Lösung für die Daten aus Aufgabe 1d) und $\alpha = 0.5$.