Technische Universität Dortmund
Fakultät Statistik
Lehrstuhl Computergestützte Statistik

Uwe Ligges
Marieke Stolte
Luca Sauer
Rudi Zulauf

Übung zur Vorlesung

Computergestützte Statistik

Wintersemester 2022/2023

Musterlösung zu Übungsblatt Nr. 10

Aufgabe 1 **(4 Punkte)**

1.)

$$
\begin{aligned}
f_1(\boldsymbol{x}) &= \boldsymbol{x}^T\boldsymbol{x} = (x_1 \cdots x_n)(x_1 \cdots x_n)^T = \sum_{i=1}^n x_i^2 = ||x_i||_2^2 \\
\nabla f_1 &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} \\ \vdots \\ \frac{\partial f_1}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 2x_1 \\ \vdots \\ 2x_n \end{pmatrix} = 2\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 2\boldsymbol{x} \\
\nabla^2 f_1 &= \begin{pmatrix} \frac{\partial^2 f_1}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f_1}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f_1}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f_1}{\partial x_n \partial x_n} \end{pmatrix} = \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2 \end{pmatrix} = 2I_n
\end{aligned}
$$

2.)

$$
\begin{aligned}
f_2(\boldsymbol{x}) &= \boldsymbol{x}^T\boldsymbol{D}\boldsymbol{x} = (x_1 \cdots x_n)\boldsymbol{D}(x_1 \cdots x_n)^T = (x_1 \cdots x_n)\begin{pmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nn} \end{pmatrix}(x_1 \cdots x_n)^T \\
&= (x_1 \cdots x_n)\left(\sum_{k=1}^n x_k d_{1k}, \ldots, \sum_{k=1}^n x_k d_{nk}\right)^T = \sum_{k=1}^n x_k \left(\sum_{j=1}^n x_j d_{jk}\right) = \sum_{k=1}^n \sum_{j=1}^n x_k x_j d_{jk} \\
\nabla f_2 &= \begin{pmatrix} \frac{\partial f_2}{\partial x_1} \\ \vdots \\ \frac{\partial f_2}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 2x_1 d_{11} + 2x_2 d_{12} + \cdots + 2x_n d_{1n} \\ \vdots \\ 2x_1 d_{n1} + 2x_2 d_{n2} + \cdots + 2x_n d_{nn} \end{pmatrix} = \begin{pmatrix} 2\sum_{k=1}^n x_k d_{1k} \\ \vdots \\ 2\sum_{k=1}^n x_k d_{kn} \end{pmatrix} = 2\boldsymbol{D}\boldsymbol{x} \\
\nabla^2 f_2 &= \begin{pmatrix} \frac{\partial^2 f_2}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f_2}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f_2}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f_2}{\partial x_n \partial x_n} \end{pmatrix} = 2\begin{pmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{1n} & \cdots & d_{nn} \end{pmatrix} = 2\boldsymbol{D}
\end{aligned}
$$

3.)

$$f_3(\boldsymbol{x}) = \sum_{i=1}^{n} \boldsymbol{x}^T \boldsymbol{D}_i \boldsymbol{x} = \sum_{i=1}^{n} (x_1 \cdots x_n) \boldsymbol{D}_i (x_1 \cdots x_n)^T = \sum_{i=1}^{n} (x_1 \cdots x_n) \begin{pmatrix} d_{i11} & \cdots & d_{i1n} \\ \vdots & \ddots & \vdots \\ d_{in1} & \cdots & d_{inn} \end{pmatrix} (x_1 \cdots x_n)^T$$

$$= \sum_{i=1}^{n} (x_1 \cdots x_n) \left(\sum_{k=1}^{n} x_k d_{i1k}, \ldots, \sum_{k=1}^{n} x_k d_{ink} \right)^T = \sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{j=1}^{n} x_k x_j d_{ijk}$$

$$\nabla f_3 = \begin{pmatrix} \frac{\partial f_3}{\partial x_1} \\ \vdots \\ \frac{\partial f_3}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} 2x_1 d_{i11} + 2x_2 d_{i12} + \cdots + 2x_n d_{i1n} \\ \vdots \\ \sum_{i=1}^{n} 2x_1 d_{in1} + 2x_2 d_{n2} + \cdots + 2x_n d_{inn} \end{pmatrix} = \begin{pmatrix} 2\sum_{i=1}^{n} \sum_{k=1}^{n} x_k d_{i1k} \\ \vdots \\ 2\sum_{i=1}^{n} \sum_{k=1}^{n} x_k d_{ikn} \end{pmatrix}$$

$$= \sum_{i=1}^{n} 2\boldsymbol{D}_i \boldsymbol{x}$$

$$\nabla^2 f_3 = \begin{pmatrix} \frac{\partial^2 f_3}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f_3}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f_3}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f_3}{\partial x_n \partial x_n} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} 2d_{i11} & \sum_{i=1}^{n} 2d_{i12} & \cdots & \sum_{i=1}^{n} 2d_{i1n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} 2d_{i1n} & \sum_{i=1}^{n} 2d_{i2n} & \cdots & \sum_{i=1}^{n} 2d_{inn} \end{pmatrix} = \sum_{i=1}^{n} 2\boldsymbol{D}_i$$

4.)

$$f_4(\boldsymbol{x}) = -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) = -\exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right)$$

$$\nabla f_4 = \begin{pmatrix} \frac{\partial f_4}{\partial x_1} \\ \vdots \\ \frac{\partial f_4}{\partial x_n} \end{pmatrix} = \begin{pmatrix} x_1 \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \\ \vdots \\ x_n \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \end{pmatrix} = \boldsymbol{x} \cdot \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right)$$

$$\nabla^2 f_4 = \begin{pmatrix} \frac{\partial^2 f_4}{\partial x_1 \partial x_4} & \cdots & \frac{\partial^2 f_4}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f_4}{\partial x_n \partial x_4} & \cdots & \frac{\partial^2 f_4}{\partial x_n \partial x_n} \end{pmatrix}$$

$$= \begin{pmatrix} -x_1^2 \cdot \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) + \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) & \cdots & -x_1 x_n \cdot \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \\ -x_1 x_2 \cdot \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) & \ddots & -x_2 x_n \cdot \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \\ \vdots & \cdots & \vdots \\ -x_1 x_n \cdot \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) & \cdots & -x_n^2 \cdot \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) + \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \end{pmatrix}$$

$$= \begin{pmatrix} -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right)(x_1^2 - 1) & \cdots & -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) x_1 x_n \\ -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) x_1 x_2 & \ddots & -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) x_2 x_n \\ \vdots & \cdots & \vdots \\ -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) x_1 x_n & \cdots & -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right)(x_n^2 - 1) \end{pmatrix}$$

$$= -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right) \begin{pmatrix} (x_1^2 - 1) & x_1 x_2 & \cdots & x_1 x_n \\ x_1 x_2 & (x_2^2 - 1) & \cdots & x_2 x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 x_n & x_2 x - n & \cdots & (x_n^2 - 1) \end{pmatrix} = -\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right)(\boldsymbol{x}\boldsymbol{x}^T - I_n)$$

Zeigen Sie, dass $\boldsymbol{x}^* = \mathbf{0}$ ein lokales Minimum aller Funktionen ist.

1.) $\nabla f_1 \overset{!}{=} \mathbf{0} \iff 2\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{x} = \mathbf{0}$

Daraus folgt, dass $\boldsymbol{x}^* = \mathbf{0}$ ein Kandidat für ein lokales Minimum ist.

$\nabla^2 f_1 = 2I_n$ ist positiv definit. Somit ist $\boldsymbol{x}^* = \mathbf{0}$ ein lokales (und – da f_1 quasi-konvex – sogar ein globales) Minimum.

2.) $\nabla f_2 \overset{!}{=} \mathbf{0} \iff 2\boldsymbol{D}\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{D}\boldsymbol{x} = \mathbf{0}$.

Da $\boldsymbol{D}$ positiv definit und symmetrisch ist, ist $\boldsymbol{D}$ invertierbar. Somit gilt:
$\boldsymbol{D}\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{D}^{-1}\boldsymbol{D}\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{x} = \mathbf{0}$.

Daraus folgt, dass $\boldsymbol{x}^* = \mathbf{0}$ ein Kandidat für ein lokales Minimum ist.

$\nabla^2 f_2 = 2\boldsymbol{D}$ ist ebenfalls positiv definit. Somit ist $\boldsymbol{x}^* = \mathbf{0}$ ein lokales (und – da f_2 quasi-konvex – sogar ein globales) Minimum.

3.) $\nabla f_3(\boldsymbol{x}^*) = \nabla f_3(\mathbf{0}) = \sum_{i=1}^{n} 2\boldsymbol{D}_i\mathbf{0} = \mathbf{0}$.

$\nabla^2 f_3 = \sum_{i=1}^{n} 2\boldsymbol{D}_i$ ist positiv definit, da alle Matrizen $\boldsymbol{D}_i$ positiv definit sind. Daraus folgt, dass es sich bei $\boldsymbol{x}^* = \mathbf{0}$ um ein lokales (und – da f_3 quasi-konvex – sogar ein globales) Minimum handelt.

4.) $\nabla f_4 \overset{!}{=} \mathbf{0} \iff \boldsymbol{x} \cdot \underbrace{\exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}\right)}_{>0} = \mathbf{0} \iff \boldsymbol{x} = \mathbf{0}$

Daraus folgt, dass $\boldsymbol{x}^* = \mathbf{0}$ ein Kandidat für ein lokales Minimum ist.

Da $\nabla f_4(\boldsymbol{x}^*) = \nabla f_4(\mathbf{0}) = -\exp\left(-\frac{1}{2}\mathbf{0}^T\mathbf{0}\right)\left(\mathbf{0}\mathbf{0}^T - I_n\right) = I_n$, ist ∇f_4 im Punkt $\boldsymbol{x}^* = \mathbf{0}$ positiv definit. Somit ist $\boldsymbol{x}^* = \mathbf{0}$ ein lokales (und – da f_4 quasi-konvex – sogar ein globales) Minimum.