Universität Salzburg | Florian Graf

Machine Learning

Übungsblatt **3** | 20 Punkte

Aufgabe 1. *Mehrdimensionale Normalverteilung – Teil II* 16 P.

Die d-dimensionale standard Normalverteilung $\mathcal{N}(\mathbf{0},\mathbf{1})$ ist die Wahrscheinlichkeitsverteilung für Zufallsvektoren $\mathbf{X} = (X_1, \dots X_d)$, wobei alle Komponenten voneinander unabhängig sind und einer eindimensionalen standard Normalverteilung folgen, d.h., $X_i \sim \mathcal{N}(0,1)$.

(a) Bestimmen Sie die Dichtefunktion $f(x_1, \dots, x_d)$ von $\mathbf{X}$.
Erinnerung: Es gilt $\mathbb{P}[X_1 < t_1, \dots, X_d < t_d] = \int_{-\infty}^{t_d} \cdots \int_{-\infty}^{t_1} f(x_1, \dots, x_d)\, dx_1 \dots dx_d$.

(b) Bestimmen Sie die Niveaumenge der Dichtefunktion zum Niveau $(2\pi)^{-d/2}\exp(1/2)$, d.h., die Menge $\{(x_1, \dots, x_d) \in \mathbb{R}^d : f(x_1, \dots, x_d) = (2\pi)^{-d/2}\exp(-1/2)\}$. Welchem geometrischen Objekt entspricht sie?

Es sei nun $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ mit $\boldsymbol{\mu} \in \mathbb{R}^d$. Erinnerung. Wir können annehmen, dass $\mathbf{Y} = \mathbf{AX} + \boldsymbol{\mu}$ mit $\mathbf{AA}^\top = \boldsymbol{\Sigma}$. Außerdem nehmen wir an, dass $\ker(\boldsymbol{\Sigma}) = \{\mathbf{0}\}$ (d.h. falls $\boldsymbol{\Sigma}\mathbf{x} = \mathbf{0}$, dann ist $\mathbf{x} = \mathbf{0}$).
In diesem Fall ist die Dichte von $\mathbf{Y}$ gegeben durch

$$g(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right) .$$

(c) Zeigen Sie, dass $\boldsymbol{\Sigma}$ positiv definit ist, also, dass für alle $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^d$ gilt $\mathbf{v}^\top\boldsymbol{\Sigma}\mathbf{v} > 0$.

Wir betrachten nun die Niveaumengen der Dichtefunktion von $\mathbf{Y}$. Da $\boldsymbol{\Sigma}$ positiv definit ist existiert eine orthogonale Matrix $\mathbf{O}$ und eine diagonale Matrix $\boldsymbol{\Lambda}$ so, dass $\boldsymbol{\Sigma} = \mathbf{O}\boldsymbol{\Lambda}\mathbf{O}^\top$.

(d) Drücken Sie $\boldsymbol{\Sigma}$ durch die Spaltenvektoren $\mathbf{v}_i$ der Matrix $\mathbf{O} = (\mathbf{v}_1 \dots \mathbf{v}_d)$ und der Diagonaleinträge der Matrix $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_d)$ aus. Begründen Sie außerdem, wieso alle λ_i positiv sind.

(e) Bestimmen Sie ausgehend von dieser Darstellung die inverse Matrix $\boldsymbol{\Sigma}^{-1}$.
Hinweis. Da $\mathbf{O}$ orthogonal ist, ist $\mathbf{v}_1^\top\mathbf{v}_1 = \mathbf{v}_2^\top\mathbf{v}_2 = 1$ und $\mathbf{v}_1^\top\mathbf{v}_2 = 0$.

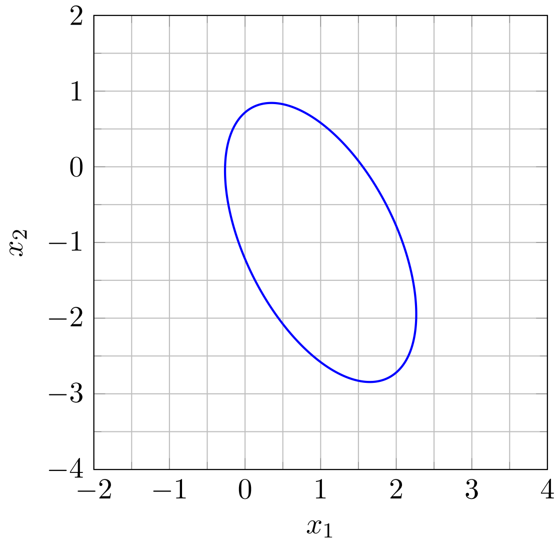
(f) Bestimmen Sie für $c > 0$ die Menge $\{\mathbf{y} \in \mathbb{R}^d\colon (\mathbf{y}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu}) = c\}$. Wählen Sie dazu den Ansatz $\mathbf{y} - \boldsymbol{\mu} = \sum_{i=1}^d a_i\mathbf{v}_i$.

(g) Wie lassen sich für einen gegebenen Vektor $\mathbf{y} \in \mathbb{R}^d$ die Koeffizienten a_i bestimmen?

Es sei nun $d = 2$, $\boldsymbol{\Sigma} = \frac{1}{2}\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ und $\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(h) Bestimmen Sie die Niveaumenge $\{\mathbf{y} \in \mathbb{R}^2 : g(\mathbf{y}) = \frac{e^{-\frac{1}{2}}}{4\pi}\}$. Zeichnen Sie die Niveaumenge handschriftlich in ein Koordinatensystem ein, beschreiben Sie ihre Form und markieren Sie wesentliche Punkte und Vektoren.

(i) Ergänzen Sie ihre Zeichnung mit einer (typischen) Stichprobe mit Umfang ≥ 20 der Zufallsvariable $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.



Die folgende Abbildung zeigt die Niveaumenge der Dichtefunktion einer Zufallsvariable $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ zum Niveau $\frac{1}{\pi}\exp(-\frac{1}{2})$.

(j) Bestimmen Sie $\boldsymbol{\mu}$ und $\mathbf{\Sigma}$ unter der Annahme, dass $\det(\mathbf{\Sigma}) = 4$.
Hinweis. Wenn Sie das Übungsblatt (in Originalgröße) ausdrucken haben die Gitterlinien Abstand 1cm und Sie können Entfernungen abmessen.

Aufgabe 2. *Mehrdimensionale Normalverteilung* 4 P.

Es seien $Y_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ und $Y_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ zwei voneinander unabhängige univariat-normalverteilte Zufallsvariablen. Bestimmen Sie die Verteilung von $Z = a_1 Y_1 + a_2 Y_2$. Gehen Sie dabei folgendermaßer vor.

(a) Welcher Verteilung folgt $\mathbf{Y} = \begin{pmatrix} Y_1 & Y_2 \end{pmatrix}^\top$?

(b) Wie lässt sich Z direkt durch den *Vektor* $\mathbf{Y}$ ausdrücken?

(c) Schließen Sie auf die Verteilung von Z.

Erinnerung. $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$ bedeutet, dass $\mathbf{Y} = \mathbf{A}\mathbf{X} + \boldsymbol{\mu}$ mit $\mathbf{A}\mathbf{A}^\top = \mathbf{\Sigma}$ mit $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{1})$, $\boldsymbol{\mu} \in \mathbb{R}^d$ und $\mathbf{A} \in \mathbb{R}^{d \times k}$.