Übungen zur Vorlesung „Hidden Markov Models“

Blatt 12

Abgabetermin: Dienstag, 23.01.2024, bis 14:15 Uhr in der Vorlesung.
(Bitte nur maximal zu zweit abgeben.)

Aufgabe 1 (4 Punkte)

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und sei $\mathcal{G} \subseteq \mathcal{F}$ eine Sub-σ-Algebra. Weiter sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable und $\phi \colon \mathbb{R} \to \mathbb{R}$ konvex. Wir nehmen an, dass X und $\phi(X)$ integrierbar sind.

(i) Zeigen Sie die bedingte Version der Jensen'schen Ungleichung, das heißt

$$\phi\big(\mathbb{E}[X|\mathcal{G}]\big) \le \mathbb{E}[\phi(X)|\mathcal{G}] \tag{1}$$

durch Reduktion auf den unbedingten Fall. Verwenden Sie hierzu (ohne Beweis), dass eine reguläre Version der bedingten Verteilung von X gegeben $\mathcal{G}$ existiert.

(ii) Untersuchen Sie, wann Gleichheit in (1) gilt.

(iii) Sei Q ein weiteres Wahrscheinlichkeitsmaß auf dem Messraum $(\Omega, \mathcal{F})$ mit $P \ll Q$. Zeigen Sie, dass die Kullback-Leibler-Divergenz von P und Q, gegeben durch

$$\mathbb{E}_P\Big[\log\Big(\frac{dP}{dQ}\Big)\Big],$$

nicht-negativ ist.

Aufgabe 2 (8 Punkte)

Wir betrachten ein HMM mit $X = \{1, \ldots, d\}$, $Y = \mathbb{R}$. Wir nehmen an, dass

$$Y_k = m(X_k) + \sqrt{v(X_k)}\eta_k, \quad \eta_k \text{ iid } \sim \mathcal{N}(0,1),$$

für zwei Funktionen $m \colon X \to \mathbb{R}$, $v \colon X \to \mathbb{R}_{++}$. Hier setzen wir $\mathbb{R}_{++} := (0, \infty)$. Wir definieren

$$\Theta := \Sigma_d \times \Delta_d \times \mathbb{R}^d \times \mathbb{R}^d_{++},$$

wobei Σ_d die Menge aller $d \times d$ stochastischen Matrizen mit strikt positiven Einträgen und Δ_d die Menge aller d-dimensionalen stochastischen Vektoren mit strikt positiven Einträgen bezeichnet. Ziel ist ist, den gesamten Parameter $\theta \in \Theta$ zu schätzen. Hierbei betrachten wir die Funktionen $m = (m_1, \ldots, m_d) \in \mathbb{R}^d$ und $v = (v_1, \ldots, v_d) \in \mathbb{R}^d_{++}$ als Vektoren.

(i) Begründen Sie kurz, warum die EM-Annahmen für $(Q^\theta)_{\theta\in\Theta}$ und $(G^\theta)_{\theta\in\Theta}$ gelten.

(ii) Zeigen Sie, dass

$$\begin{aligned} R_N(\theta, \theta') = &-\sum_{k=0}^{N}\sum_{l=1}^{d}\Big(\frac{(y_k - m_l)^2}{2v_l} + \log(\sqrt{v_l})\Big)(\pi^{\theta'}_{k|N})_l \\ &+ \sum_{k=1}^{N}\sum_{l,l'=1}^{d}\log(Q(l,\{l'\}))(\pi^{\theta'}_{k-1,k|N})_{ll'} + \sum_{l=1}^{d}\log(\nu_l)(\pi^{\theta'}_{0|N})_l \\ &- \text{ein Term der unabhängig von } \theta \text{ ist.} \end{aligned}$$

(iii) Zeigen Sie, dass für

$$\begin{aligned}
Q(i,\{j\}) &:= \frac{\sum_{k=1}^N (\pi_{k-1,k|N}^{\theta'})_{ij}}{\sum_{k=1}^N (\pi_{k-1|N}^{\theta'})_i}, \\
\nu_i &:= (\pi_{0|N}^{\theta'})_i, \\
m_i &:= \frac{\sum_{k=0}^N y_k (\pi_{k|N}^{\theta'})_i}{\sum_{k=0}^N (\pi_{k|N}^{\theta'})_i}, \\
v_i &:= \frac{\sum_{k=0}^N (y_k - m_i)^2 (\pi_{k|N}^{\theta'})_i}{\sum_{k=0}^N (\pi_{k|N}^{\theta'})_i},
\end{aligned}$$

gilt, dass $(Q,\nu,m,v) = \operatorname{argmax} R_N(\theta,\theta')$.