Hochschule RheinMain
Prof. Dr. D. Lehmann SS 2023

Lösungen zum 11. Übungsblatt
Stochastik II

Aufgabe 1) Wir benötigen die folgenden Gleichungen aus dem week11b.pdf: Wegen

$$\int_{\mathbb{R}^n} p_\theta(x)\, d^n x \;=\; 1$$

ist

$$\begin{aligned} 0 \;=\; \frac{\partial}{\partial\theta_k}\int_{\mathbb{R}^n} p_\theta(x)\, d^n x \;&=\; \int_{\mathbb{R}^n} \frac{\partial p_\theta}{\partial\theta_k}(x)\, d^n x \\ &=\; \int_{\mathbb{R}^n} \frac{\frac{\partial p_\theta}{\partial\theta_k}(x)}{p_\theta(x)}\, p_\theta(x)\, d^n x \\ &=\; \int_{\mathbb{R}^n} \frac{\partial \log p_\theta(x)}{\partial\theta_k}\, p_\theta(x)\, d^n x \;=\; \mathsf{E}\Big[\,\frac{\partial \log p_\theta}{\partial\theta_k}\,\Big] \end{aligned} \tag{1}$$

Diese Identität tun wir noch einmal nach θ_ℓ ableiten und bekommen

$$\begin{aligned} 0 \;&=\; \frac{\partial}{\partial\theta_\ell}\int_{\mathbb{R}^n} \frac{\partial \log p_\theta(x)}{\partial\theta_k}\, p_\theta(x)\, d^n x \\ &=\; \int_{\mathbb{R}^n} \frac{\partial^2 \log p_\theta(x)}{\partial\theta_\ell\partial\theta_k}\, p_\theta(x)\, d^n x \;+\; \int_{\mathbb{R}^n} \frac{\partial \log p_\theta(x)}{\partial\theta_k}\, \frac{\partial p_\theta}{\partial\theta_\ell}(x)\, d^n x \\ &=\; \int_{\mathbb{R}^n} \frac{\partial^2 \log p_\theta(x)}{\partial\theta_\ell\partial\theta_k}\, p_\theta(x)\, d^n x \;+\; \int_{\mathbb{R}^n} \frac{\partial \log p_\theta(x)}{\partial\theta_k}\, \frac{\partial \log p_\theta(x)}{\partial\theta_\ell}\, p_\theta(x)\, d^n x \end{aligned}$$

Also haben wir

$$\begin{aligned} \int_{\mathbb{R}^n} \frac{\partial \log p_\theta(x)}{\partial\theta_k}\, \frac{\partial \log p_\theta(x)}{\partial\theta_\ell}\, p_\theta(x)\, d^n x \;&=\; -\int_{\mathbb{R}^n} \frac{\partial^2 \log p_\theta(x)}{\partial\theta_\ell\partial\theta_k}\, p_\theta(x)\, d^n x \\ &=\; -\,\mathsf{E}\Big[\,\frac{\partial^2 \log p_\theta}{\partial\theta_\ell\partial\theta_k}\,\Big] \\ &=\; +\,I(\theta)_{\ell,k} \;=\; +\,I(\theta)_{k,\ell} \end{aligned} \tag{2}$$

Damit können wir jetzt die Formeln für den Erwartungswert und die Varianz verifizieren: Die allgemeine Darstellung der Minimum-Varianz-Schätzer war

$$\hat\theta(x) \;=\; \theta \;+\; I^{-1}(\theta)\,\nabla_\theta \log p_\theta(x) \tag{3}$$

Der k-te Schätzer ist also gegeben durch

$$\hat{\theta}_k(x) \;=\; \theta_k \;+\; \sum_{\ell=1}^{m} [I^{-1}(\theta)]_{k,\ell}\, \frac{\partial \log p_\theta(x)}{\partial \theta_\ell} \tag{4}$$

und wir bekommen

$$\mathsf{E}[\,\hat{\theta}_k\,] \;=\; \theta_k \;+\; \sum_{\ell=1}^{m} [I^{-1}(\theta)]_{k,\ell}\; \underbrace{\mathsf{E}\Big[\,\frac{\partial \log p_\theta}{\partial \theta_\ell}\,\Big]}_{\overset{(1)}{=}\,0} \;=\; \theta_k \tag{5}$$

Damit ist der Teil (a) bewiesen. Berechnen wir die Varianz. Wir haben

$$\begin{aligned}
\mathsf{V}[\,\hat{\theta}_k\,] \;&=\; \mathsf{E}\big[\,(\hat{\theta}_k - \mathsf{E}[\hat{\theta}_k])^2\,\big] \\
&\overset{(a)}{=}\; \mathsf{E}\big[\,(\hat{\theta}_k - \theta_k)^2\,\big] \\
&\overset{(4)}{=}\; \mathsf{E}\Big[\,\Big(\sum_{\ell=1}^{m}[I^{-1}(\theta)]_{k,\ell}\,\frac{\partial \log p_\theta(x)}{\partial \theta_\ell}\Big)^2\,\Big] \\
&=\; \sum_{j,\ell=1}^{m}[I^{-1}(\theta)]_{k,j}\,[I^{-1}(\theta)]_{k,\ell} \,\times\, \mathsf{E}\Big[\,\frac{\partial \log p_\theta(x)}{\partial \theta_j}\;\frac{\partial \log p_\theta(x)}{\partial \theta_\ell}\,\Big] \\
&\overset{(2)}{=}\; \sum_{j,\ell=1}^{m}[I^{-1}(\theta)]_{k,j}\,[I^{-1}(\theta)]_{k,\ell} \,\times\, [I(\theta)]_{j,\ell} \\
&=\; \sum_{j,\ell=1}^{m}[I^{-1}(\theta)]_{k,j} \,\times\, [I(\theta)]_{j,\ell} \,\times\, [I^{-1}(\theta)]_{\ell,k} \\
&=\; \big[\,I^{-1}(\theta)\;I(\theta)\;I^{-1}(\theta)\,\big]_{k,k} \;=\; [I^{-1}(\theta)]_{k,k}
\end{aligned} \tag{6}$$

und damit ist auch der Teil (b) bewiesen.

Aufgabe 2) Für das Standardbeispiel 1 mit unabhängigen, normalverteilten Zufallszahlen haben wir die Fisher-Informationsmatrix $I = I(\mu,\nu)$ mit $\nu := \sigma^2$ bereits in dem week11a.pdf berechnet, auf der letzten Seite. Das Resultat war

$$I(\theta) \;=\; I(\mu,\nu) \;=\; \begin{pmatrix} \frac{n}{\nu} & 0 \\ 0 & \frac{n}{2\nu^2} \end{pmatrix}$$

und damit

$$I^{-1}(\mu,\nu) \;=\; \begin{pmatrix} \frac{\nu}{n} & 0 \\ 0 & \frac{2\nu^2}{n} \end{pmatrix}$$

Den Gradienten von $\log p_\theta = \log p_{\mu,\nu}$ hatten wir ebenfalls schon berechnet, mit

$$\begin{aligned}
p_\theta(x) \;&=\; p_{\mu,\nu}(x_1,\cdots,x_n) \;=\; \prod_{i=1}^{n}\Big\{\,\tfrac{1}{\sqrt{2\pi\nu}}\,e^{-\frac{(x_i-\mu)^2}{2\nu}}\,\Big\} \\
\log p_\theta(x) \;&=\; \log p_{\mu,\nu}(x_1,\cdots,x_n) \;=\; -\tfrac{n}{2}\log(2\pi\nu) \;-\; \sum_{i=1}^{n}\tfrac{(x_i-\mu)^2}{2\nu}
\end{aligned}$$

bekamen wir die folgenden Ableitungen:

$$
\begin{aligned}
\frac{\partial \log p_{\mu,\nu}}{\partial \mu} &= \sum_{i=1}^{n} \frac{x_i-\mu}{\nu} \\
\frac{\partial \log p_{\mu,\nu}}{\partial \nu} &= -\frac{n}{2\nu} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\nu^2}
\end{aligned}
$$

Aus der allgemeinen Gleichung

$$\hat{\theta}(x) = \theta + I^{-1}(\theta)\, \nabla_\theta \log p_\theta(x)$$

folgt dann für diesen konkreten Fall die Darstellung

$$\begin{pmatrix} \hat{\mu} \\ \hat{\nu} \end{pmatrix} = \begin{pmatrix} \mu \\ \nu \end{pmatrix} + \begin{pmatrix} \frac{\nu}{n} & 0 \\ 0 & \frac{2\nu^2}{n} \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} \frac{x_i-\mu}{\nu} \\ -\frac{n}{2\nu} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\nu^2} \end{pmatrix}$$

Also,

$$\hat{\mu} = \mu + \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

und

$$\hat{\nu} = \nu - \nu + \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Damit sind die Formeln aus (a) und (b) bewiesen.