ORF524 - Problem Set 2

Bachir EL KHADIR

October 10, 2015

Question 1

Question 2

Let's first note that:

$$l(\theta) = \frac{1}{\int_{\mathbb{R}^d} h(x) e^{\alpha(\theta)^T T(x)} \mathrm{dx}} = l(\alpha(\theta))$$

As a result, f^θ is determined entirely by $\alpha(\theta)$, we can then denote it $f_{\alpha(\theta)}$ As a result

$$P = \{f_\alpha | \alpha \in \alpha(\Theta)\}$$

Question 3

Question 4

$$\begin{aligned}
\mathcal{N}^n_{\mu,\mu}(x) &= \frac{1}{(\sqrt{2\pi\mu})^n} e^{-\sum_i \frac{(x_i-\mu)^2}{2\mu}} && (1)\\
&= \frac{1}{(\sqrt{2\pi\mu})^n} e^{-\frac{\sum_i x_i^2}{2\mu} - \sum x_i - n\frac{\mu}{2}} && (2)\\
&= \frac{1}{(\sqrt{2\pi\mu})^n} e^{-\frac{\sum_i x_i^2}{2\mu} - n\frac{\mu}{2}} e^{-\sum x_i} && (3)\\
&= g_\mu(\sum x_i^2) f(x) && (4)
\end{aligned}$$

$$T(x) = \sum x_i^2$$

Let $x, x' \in \mathbb{R}^d$, the quantity

$$\begin{aligned}
\frac{\mathcal{N}^n_{\mu,\mu}(x)}{\mathcal{N}^n_{\mu,\mu}(x')} &= e^{-\frac{1}{2\mu}(T(x)-T(x'))} \frac{f(x)}{f(x')} && (5)\\
& && (6)
\end{aligned}$$

is independant of μ if only if $T(x) = T(x')$, therefore T is minimal sufficient.

For $n = 1$, $T = T_0^2$, so T_0 is sufficient. It is no minimal because

$$\begin{aligned}
\frac{\mathcal{N}_{\mu,\mu}(1)}{\mathcal{N}_{\mu,\mu}(-1)} &= e^{-\frac{1}{2\mu}(T(1)-T(-1))} \frac{f(1)}{f(-1)} && (7)\\
&= \frac{f(1)}{f(-1)} && (8)
\end{aligned}$$

is independant of μ, but $T_0(1) \neq T_0(-1)$.

Question 5

For n

Question 6

The log-likelihood function:

$$\begin{aligned}
\mathcal{L}(\theta; x) &= \log(\Pi_i f(x_i|\theta)) && \text{because iid} && (9)\\
&= \log \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\sum_i \frac{(x_i-\mu)^2}{2\sigma^2}} && && (10)\\
&= -n\log(\sqrt{2\pi}) - n\log\sigma - \sum_i \frac{(x_i-\mu)^2}{2\sigma^2} && && (11)\\
& && && (12)
\end{aligned}$$

$$\frac{d\mathcal{L}}{d\sigma} = -\frac{n}{\sigma} + \frac{\sum_i (x_i-\mu)^2}{\sigma^3}$$

$$\frac{d\mathcal{L}}{d\mu} = \frac{\sum_i x_i - \mu}{\sigma^2} = \frac{\bar{x}-\mu}{\sigma^2}$$

MLE

$$\theta = (\bar{x}, \frac{1}{n}\sum_i (x_i - \bar{x})^2)$$

Question 7

$$\theta = (p_l, \mu_l, \Sigma_l)_l$$

$$\begin{aligned}
Q(\theta, \theta') &= \mathbb{E}^{\theta'}[\log \mathcal{L}^n(X, L|\theta)|X] && (13)\\
&= \mathbb{E}^{\theta'}[\log \Pi_i \mathcal{L}(X_i, L_i; \theta)|X] && (14)\\
&= \sum_i \mathbb{E}^{\theta'}[\log \mathcal{L}(X_i, L_i; \theta)|X_i] && (15)\\
& && (16)
\end{aligned}$$

$$\mathbb{E}^{\theta'}[\log \mathcal{L}(X_i, L_i; \theta)|X_i] = \sum_l \mathbb{P}(L = l|X_i; \theta') \log \mathcal{L}(X_i, l|\theta) \tag{17}$$

$$\mathbb{P}(L = l|X_i; \theta') = \frac{f(L = l, X = X_i|\theta')}{f(X = X_i; \theta')} = \frac{\mathbb{P}(L = l|\theta') f(X = X_i|L = l; \theta')}{\sum_k \mathbb{P}(L = k|\theta') f(X = X_i|L = k; \theta')}$$

$$\log \mathcal{L}(X_i, l|\theta) = \log$$

Question 8

- $$\begin{aligned}
\mathbb{P}(\hat{\theta} \leq x) &= \mathbb{P}\{\max_i x_i \leq x\} && (18)\\
&= \mathbb{P}(\cap_i \{x_i \leq x\}) && (19)\\
&= \Pi_i \mathbb{P}(x_i \leq x) && (20)\\
&= \min\left(1, \left(\frac{x}{\theta}\right)^n\right) && (21)\\
&= \int_{\mathbb{R}} n\frac{y^{n-1}}{\theta^n} 1_{0\leq y\leq \theta} 1_{y \leq x} \mathrm{dy} && (22)\\
&= \int^x f(y)\mathrm{dy} && (23)
\end{aligned}$$

- $$\begin{aligned}
\mathbb{E}[\hat{\theta}] &= \int_0^\theta yn\frac{y^{n-1}}{\theta^n}\mathrm{dy} && (24)\\
&= \frac{n}{n+1}\theta \neq \theta \text{ if } \theta \neq 0 && (25)
\end{aligned}$$

Question 9

$$\mathcal{L}(\theta; x) = \mathcal{L}(\theta; x_1|x_2, ...)\mathcal{L}(\theta; x_2|x_3, ...)...\mathcal{L}(\theta; x_n)$$
$$\mathbb{E}\log\mathcal{L}(x;\theta) = \sum_i \mathbb{E}\log\mathcal{L}(x_i;\theta|x_{i+1}...x_n) = -\sum_i H(x_i|x_{i+1}...x_n)$$

$$\begin{aligned}
H(X) - H(X|Y) &= \mathbb{E} log(f(Y)/f(X,Y)) && (26)\\
&\leq \log\mathbb{E}\frac{f(Y)}{f(X,Y)} && (27)\\
&= \log\int\frac{f(Y)}{f(X,Y)}f(X,Y) && (28)\\
&= \log 1 = 0 && (29)
\end{aligned}$$

Question 10

$$g(\beta) = \sum(y_i - x_i^T\beta)^2$$
$$f(\beta) = \sum(y_i - x_i^T\beta)^2) + \lambda||\beta||^2 + g(\beta) = \lambda||\beta||^2$$

$$\begin{aligned}
\nabla_\beta f &= \sum_i -2(y_i - x_i^T\beta)x_i + 2\lambda\beta && (30)\\
&= 2(\lambda\beta - \sum_i (y_i - x_i^T\beta)x_i) && (31)\\
&= 2((\lambda I_n + \sum_i x_i x_i^T)\beta + \sum_i y_i x_i) && (32)
\end{aligned}$$

The hessian of f is $F := 2(\lambda I_n + \sum_i x_i x_i^T)$. F is symetric and its eigen values are those of $\sum_i x_i x_i^T$ offset by λ. For λ large enough ($\lambda > ||\sum_i x_i x_i^T||_\infty$), the eigen values of F are all positive, and therefore f is strictly convexe and admit at most one global minimum.

In addition, there is a solution iff $\nabla f = 0$ has a solution, and the solution happens to be the minimum. Which is the case for

$$\beta = \frac{1}{2} F^{-1} \sum y_i x_i = (\lambda I_n + \sum_i x_i x_i^T)^{-1} \sum_i y_i x_i$$

Question 11

Question 12

Let $\hat{X} = (X^l)_{l \in \mathbb{N}^p : |l| \leq k}$

$$Y = poly(X) + \epsilon = \beta \hat{X} + \epsilon$$

$\beta = (\sum_i \hat{x}_i \hat{x}_i^T)^{-1} \sum_i y_i \hat{x}_i$