Stat 461-561: Quiz 1

Monday 28th January 2008

• **Exercise 1**. Let $X_i \overset{\text{i.i.d.}}{\sim} f(x|\theta^*)$ where $\theta^* \in \Theta \subseteq \mathbb{R}$. Let θ_n be the Maximum Likelihood Estimate (MLE) for n observations; that is

$$\theta_n = \arg\max_{\theta \in \Theta} \sum_{i=1}^{n} \log f(X_i|\theta)$$

Under 'suitable' regularity assumptions, we have

$$\sqrt{n}(\theta_n - \theta^*) \xrightarrow{\text{D}} \mathcal{N}(0, \sigma^2)$$

[1 point] *Establish* the expression of σ^2.

See lecture notes
[1 point] Propose an asymptotically consistent estimate of σ^2?

We would use the following consistent estimate.

$$\widehat{\sigma^2} = \left(-\frac{1}{n} \sum_{i=1}^{n} \left. \frac{\partial^2 \log f(X_i|\theta)}{\partial \theta^2} \right|_{\theta_n} \right)^{-1}$$

• **Exercise 2**. Suppose $\{X_k\}_{k \geq 1}$ be a Markov process such that $X_1 = x_1$ is fixed and for $k \geq 2$ by the equation

$$X_k = \alpha X_{k-1} + \sigma V_k$$

where $V_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ and $|\alpha| < 1$.
[1 point] Let $\theta = (\alpha, \sigma^2)$. Write down the expression of the log-likelihood $l(\theta)$ of the observations $x_{1:n} := (x_1, x_2, ..., x_n)$.

We have for the likelihood

$$\begin{aligned} L(\theta) &= \prod_{k=2}^{n} \mathcal{N}(x_k; \alpha x_{k-1}, \sigma^2) \\ &= \prod_{k=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_k - \alpha x_{k-1})^2}{2\sigma^2} \right) \end{aligned}$$

so

$$l(\theta) = -\frac{n-1}{2} \log(2\pi\sigma^2) - \sum_{k=2}^{n} \frac{(x_k - \alpha x_{k-1})^2}{2\sigma^2}.$$

[2 points] Compute the Maximum Likelihood Estimate (MLE) $\widehat{\theta}$ of θ.

We have

$$\begin{aligned}
\frac{\partial l(\theta)}{\partial \alpha} &= \frac{\sum_{k=2}^{n} x_{k-1}\left(x_k - \alpha x_{k-1}\right)}{\sigma^2}, \\
\frac{\partial l(\theta)}{\partial \sigma^2} &= -\frac{n-1}{2\sigma^2} + \sum_{k=2}^{n} \frac{\left(x_k - \alpha x_{k-1}\right)^2}{2\sigma^4}
\end{aligned}$$

so we obtain

$$\begin{aligned}
\widehat{\alpha} &= \frac{\sum_{k=2}^{n} x_{k-1} x_k}{\sum_{k=2}^{n} x_{k-1}^2}, \\
\widehat{\sigma^2} &= \frac{\sum_{k=2}^{n} \left(x_k - \widehat{\alpha} x_{k-1}\right)^2}{n-1}.
\end{aligned}$$

[1 point] Assume from now on that $X_1 \sim \mathcal{N}\left(0, \frac{\sigma^2}{1-\alpha^2}\right)$. Can you compute the MLE analytically? If not, propose an iterative numerical method to find the MLE. How would you initialize it?

In this case, there is no closed-form expression for the MLE but we can use a Newton-Raphson method. It would be sensible to initialize this algorithm at the estimate obtained if X_1 was assumed non-random (i.e. the estimate discussed in the previous question)

[1 point] Consider the following pseudo log-likelihood function

$$\widetilde{l}_n(\theta) = \sum_{k=1}^{n-1} \log p\left(x_k, x_{k+1} \mid \theta\right)$$

where $p\left(x_k, x_{k+1} \mid \theta\right)$ is the marginal distribution $\left(X_k, X_{k+1}\right)$. Establish the expression of $\widetilde{l}_n(\theta)$.

We can show easily that $p\left(x_k \mid \theta\right) = \mathcal{N}\left(x_k; 0, \frac{\sigma^2}{1-\alpha^2}\right)$ (see lecture notes) so

$$\begin{aligned}
\widetilde{l}_n(\theta) &= \sum_{k=1}^{n-1} \log\left(\mathcal{N}\left(x_k; 0, \frac{\sigma^2}{1-\alpha^2}\right) \mathcal{N}\left(x_{k+1}; \alpha x_k, \sigma^2\right)\right) \\
&= \sum_{k=1}^{n-1} -\log(2\pi) + \frac{1}{2}\log\left(1-\alpha^2\right) - \log\left(\sigma^2\right) - \frac{\left(1-\alpha^2\right) x_k^2}{2\sigma^2} - \frac{\left(x_{k+1} - \alpha x_k\right)^2}{2\sigma^2} \\
&= -(n-1)\log(2\pi) + \frac{n-1}{2}\log\left(1-\alpha^2\right) - (n-1)\log\left(\sigma^2\right) \\
&\quad - \sum_{k=1}^{n-1} \frac{\left(1-\alpha^2\right) x_k^2}{2\sigma^2} - \sum_{k=1}^{n-1} \frac{\left(x_{k+1} - \alpha x_k\right)^2}{2\sigma^2}
\end{aligned}$$

[2 points] Let θ_n be the pseudo-likelihood estimate defined by

$$\theta_n = \arg\max_{\theta \in \Theta} \widetilde{l}_n(\theta)$$

Without establishing the expression of θ_n, explain why this estimate is asymptotically consistent.

We have thanks to the law of large numbers described in *Hint*

$$\lim_{n \to \infty} \frac{1}{n} \widetilde{l}_n(\theta) = \int \log p\left(x, x' \middle| \theta\right) . p\left(x, x' \middle| \theta^*\right) dx dx'.$$

which is maximized in $\theta = \theta^*$ as it is equivalent to minimize the KL divergence between $p(x, x' | \theta)$ and $p(x, x' | \theta^*)$.

[1 point] Consider now the alternative pseudo log-likelihood function

$$\bar{l}_n(\theta) = \sum_{k=1}^{n} \log p(x_k | \theta).$$

Explain why this pseudo log-likelihood function does *not* allow us to find an asymptotically consistent estimate.

In the marginal distributions, we only have information about $\left(\alpha, \sigma^2\right)$ through $\frac{\sigma^2}{1-\alpha^2}$ so it is impossible to estimate the two parameters.