

## Statistical Modeling and Inference – Problem Set #6

GROUP 3: NITI MISHRA · MIQUEL TORRENS · BÁLINT VÁN

November 30<sup>th</sup>, 2015

Solution to proposed exercises.

### Exercise 1

#### Part (a)

Given  $t_n \sim \text{Laplace}(\mu, \Lambda)$ , the likelihood function is:

$$\mathcal{L}(\text{Laplace}(t_n|\mu, \Lambda)) = \prod_{n=1}^N f(t_n|\mu, \Lambda) = \left(\frac{\Lambda}{2}\right)^N \exp\left\{-\Lambda \sum_{n=1}^N |t_n - \mu|\right\}$$

#### Part (b)

If we take the MLE:

$$\max_{\mu, \Lambda} \log \mathcal{L} = \max_{\mu, \Lambda} N \log \Lambda - \Lambda \sum_{n=1}^N |t_n - \mu| + C$$

With  $C$  containing the terms not dependent on  $\mu$  or  $\Lambda$ . Then:

$$\begin{aligned} \frac{\partial \log \mathcal{L}}{\partial \mu} = 0 &\Leftrightarrow -\Lambda \sum_{n=1}^N \frac{t_n - \mu}{|t_n - \mu|} = 0 \\ &\Leftrightarrow \sum_{n=1}^N \text{sgn}(t_n - \mu_{MLE}) = 0 \end{aligned} \tag{1}$$

Note that with odd  $N$ , for (1) to hold,  $(N-1)/2$  observations need to have  $t_n \geq \mu$  and the other  $(N-1)/2$  have  $t_n \leq \mu$ . Thus:

$$\mathbb{E}[t] = \mu_{MLE} = \text{median}(t_1, t_2, \dots, t_n)$$

is the expected value and also the median.

#### Part (c)

We derive  $\Lambda_{MLE}$ :

$$\begin{aligned} \frac{\partial \log \mathcal{L}}{\partial \Lambda} = 0 &\Leftrightarrow N \frac{1}{\Lambda} - \sum_{n=1}^N |t_n - \mu| = 0 \\ &\Leftrightarrow \Lambda_{MLE} = \left( \frac{1}{N} \sum_{n=1}^N |t_n - \mu| \right)^{-1} \end{aligned}$$

Part (d)

By the equivariance property<sup>1</sup> of the MLE:

$$\text{var}[t] = \Sigma_{MLE} = \frac{2}{\Lambda_{MLE}^2} = 2 \left( \frac{1}{N} \sum_{n=1}^N |t_n - \mu| \right)^2 .$$

---

<sup>1</sup>Wasserman's Theorem 10.14.

## Exercise 2

Part (a)

Given:

$$t_n \sim \mathcal{N}(t_n | \mathbf{x}_n, \mathbf{w}, \eta_n, q) = \frac{(\eta_n q)^{1/2}}{\sqrt{2\pi}} \exp \left\{ -\frac{\eta_n q}{2} (t_n - \phi(\mathbf{x}_n) \mathbf{w})^T (t_n - \phi(\mathbf{x}_n) \mathbf{w}) \right\}$$

and:

$$\eta_n \sim \text{Gam} \left( \eta_n | \frac{\nu}{2}, \frac{\nu}{2} - 1 \right) = \frac{\left(\frac{\nu}{2} - 1\right)^{\nu/2}}{\left(\frac{\nu}{2} - 1\right)!} \eta_n^{\left(\frac{\nu}{2} - 1\right)} \exp \left\{ -\left(\frac{\nu}{2} - 1\right) \eta_n \right\}$$

We know:

$$\begin{aligned} p(\eta_n | t_n, \mathbf{x}_n, \mathbf{w}, q) &\propto p(t_n | \eta_n, \mathbf{x}_n, \mathbf{w}, q) p(\eta_n) \\ &= C \eta_n^{\left(\frac{\nu}{2} - 1 + \frac{1}{2}\right)} \exp \left\{ -\left(\frac{\nu}{2} - 1\right) \eta_n - \frac{\eta_n q}{2} \mathbf{e}_n^T \mathbf{e}_n \right\} \\ &= C \eta_n^{\left(\frac{\nu+1}{2} - 1\right)} \exp \left\{ \left( -\left(\frac{\nu}{2} - 1\right) - \frac{q}{2} \mathbf{e}_n^T \mathbf{e}_n \right) \eta_n \right\} \\ &= C \eta_n^{\left(\frac{\nu+1}{2} - 1\right)} \exp \left\{ -\left( \frac{\nu + q \mathbf{e}_n^T \mathbf{e}_n}{2} - 1 \right) \eta_n \right\} \end{aligned}$$

Where  $C$  contains the multiplication of the terms of the density functions not dependent on  $\eta_n$ . Observe this is a Gamma distribution with parameters:

$$\begin{aligned} \alpha_n &= \frac{\nu + 1}{2} \\ \beta_n &= \frac{\nu + q \mathbf{e}_n^T \mathbf{e}_n}{2} - 1 \end{aligned}$$

This finishes the proof.

Part (b)

Given<sup>2</sup>:

$$Q(\theta, \theta') = \int \log(p(t, \eta | \theta)) p(\eta | t, \theta') d\eta = \mathbb{E}[\log(p(t, \eta | \theta))]$$

---

<sup>2</sup>Equations (4.3) and (4.4) in Omiros' notes on the appendix of the lecture.

We can use  $\theta = (q, \mathbf{w})$  and through maximum likelihood compute:

$$\begin{aligned}
\mathbb{E}[\log p(t, \eta|\theta)] &= \mathbb{E}[\log(p(t|\theta)p(\eta|t, \theta'))] \\
&= \mathbb{E}[\log p(t|\theta) + \log p(\eta|t, \theta')] \\
&= \mathbb{E}\left[\log\left(\prod_{n=1}^N \left(\frac{q\eta_n}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}qe_n^2\eta_n\right\}\right) + \log p(\eta|t, \theta')\right] \\
&= \mathbb{E}\left[\log\left(\prod_{n=1}^N \left(\frac{q\eta_n}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}qe_n^2\eta_n\right\}\right)\right] + \mathbb{E}[\log p(\eta|t, \theta')] \\
&= \mathbb{E}\left[\log\left(\left(\frac{q\eta_n}{2\pi}\right)^{\frac{N}{2}} \exp\left\{-\frac{1}{2}q\sum_{n=1}^N e_n^2\eta_n\right\}\right)\right] + c \\
&= \mathbb{E}\left[\frac{N}{2}\log q - \frac{q}{2}\sum_{n=1}^N (t_n - \phi(\mathbf{x}_n)\mathbf{w})^2\eta_n\right] + c \\
&= \frac{N}{2}\log q - \frac{q}{2}(\mathbf{t} - \Phi\mathbf{w})^T \mathbb{E}[\mathbf{H}] (\mathbf{t} - \Phi\mathbf{w}) + c
\end{aligned}$$

Where  $\mathbf{H}$  is a diagonal matrix with  $\mathbf{H}_{nn} = \eta_n$ , and  $c$  contains all terms not dependent on  $\theta$ . The distribution of  $\eta_n$  will not depend on  $\theta$  but on the previous  $\theta' = (q', \mathbf{w}')$ , and given that for  $t \sim \text{Gam}(\alpha, \beta)$  we have  $\mathbb{E}[t] = \alpha/\beta$ , then:

$$\mathbb{E}[\eta_n] = \frac{\frac{\nu+1}{2}}{\frac{\nu+q'(e'_n)^2}{2} - 1} = \frac{\nu+1}{\nu+q'(e'_n)^2 - 2} = \frac{\nu+1}{\nu+q'(t_n - \phi(\mathbf{x}_n)\mathbf{w}')^2 - 2}$$

Hence proved.

### **Exercise 3**

Part (a)

Answer.

Part (b)

Answer.

Part (c)

Answer.

Part (d)

Answer.