演習問題 1.11

対数尤度関数 (1.54) の μ と σ^2 に関する微分を 0 とおいて、式 (1.55) と式 (1.56) を確かめよ。

$$\ln p\,(\,\mathbf{X}\mid\mu,\sigma^2\,) = -\frac{1}{2\sigma^2}\sum_{n=1}^{N}(\,x_n-\mu\,)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(\,2\pi\,)$$

$\cdots$ (1.54)

$$\mu_{\mathrm{ML}} = \frac{1}{N}\sum_{n=1}^{N}x_n$$

$\cdots$ (1.55)

$$\sigma_{\mathrm{ML}}^2 = \frac{1}{N}\sum_{n=1}^{N}(\,x_n-\mu_{\mathrm{ML}}\,)^2$$

$\cdots$ (1.56)

[解]

まず、μ と σ^2 に関する微分を 0 とおいて、式 (1.55) を証明する。式 (1.54) を以下のように展開する。

$$\begin{aligned}\ln p\,(\,\mathbf{X}\mid\mu,\sigma^2\,) &= -\frac{1}{2\sigma^2}\sum_{n=1}^{N}(\,x_n-\mu\,)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(\,2\pi\,)\\ &= -\frac{1}{2\sigma^2}\left\{\sum_{n=1}^{N}x_n^2 - 2\mu\sum_{n=1}^{N}x_n + N\mu^2\right\} - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(\,2\pi\,)\end{aligned}$$

これを μ に関して微分を行うと、

$$\frac{d\ln p\,(\,\mathbf{X}\mid\mu,\sigma^2\,)}{d\mu} = -\frac{1}{2\sigma^2}\left\{-2\sum_{n=1}^{N}x_n + 2\,N\mu\right\} = 0$$

より、

$$\mu = \frac{1}{N}\sum_{n=1}^{N}x_n$$

となる。よって、式 (1.55) が示せた。

まず、μ と σ^2 に関する微分を 0 とおいて、式 (1.56) を証明する。式 (1.54) を以下のように整理する。

$$\begin{aligned}\ln p\,(\,\mathbf{X}\mid\mu,\sigma^2\,) &= -\frac{1}{2\sigma^2}\sum_{n=1}^{N}(\,x_n-\mu\,)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(\,2\pi\,)\\ &= -\frac{\sigma^{-2}}{2}\sum_{n=1}^{N}(\,x_n-\mu\,)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(\,2\pi\,)\\ &= -\frac{1}{2}\left\{\sum_{n=1}^{N}(\,x_n-\mu\,)^2\,\sigma^{-2} + N\ln\sigma^2 + N\ln(\,2\pi\,)\right\}\end{aligned}$$

これを、 σ^2 に関して微分を行うのだが、σ^2 に関して微分を行うと、式が煩雑となってしまうため、 $a = \sigma^2$ と置き、a に関して微分を行うこととする。 $a = \sigma^2$ と置くと、上記の式は、

$$= -\frac{1}{2}\left\{ \sum_{n=1}^{N} (x_n - \mu)^2 a^{-1} + N \ln a + N \ln(2\pi) \right\}$$

となり、ここで、上記の式を a に関して微分を行うと、

ここで、$\ln x^a$ の x に関する微分は、 $$(\ln x^a)' = \frac{1}{x^a} \cdot (x^a)' = \frac{a x^{a-1}}{x^a}$$ であることに注意する。

$$\frac{d \ln p(\mathbf{X} \mid \mu, \sigma^2)}{da} = -\frac{1}{2}\left\{ -\sum_{n=1}^{N} (x_n - \mu)^2 a^{-2} + N \cdot \frac{1}{a} \right\}$$

$$= -\frac{1}{2}\left\{ -\sum_{n=1}^{N} (x_n - \mu)^2 a^{-2} + N a^{-1} \right\} = 0$$

であるから、

$$\sum_{n=1}^{N} (x_n - \mu)^2 a^{-2} - N a^{-1} = 0$$

$$\sum_{n=1}^{N} (x_n - \mu)^2 - Na = 0$$

$$a = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^2$$

と整理できる。$a = \sigma^2$ であるので、

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^2$$

と書き直す。よって、式 (1.56) が示せた。