演習問題 1.30

2 つのガウス分布 $p(x) = N(x \mid \mu, \sigma^2)$ と $q(x) = N(x \mid m, s^2)$ の間のカルバックーライブラーダイバージェンス (1.113) を計算せよ。

[期待値]

ある関数 $f(x)$ の確率分布 $p(x)$ の下での平均値のこと。

$$\mathrm{E}[f] = \int p(x) f(x)\, dx$$

… (1.34)

[ガウス分布]

$$\boldsymbol{N}(x \mid \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

… (1.46)

[ガウス分布の平均]

$$\mathrm{E}[x] = \int_{-\infty}^{\infty} \boldsymbol{N}(x \mid \mu, \sigma^2)\, x\, dx = \mu$$

… (1.49)

$$\mathrm{E}[x^2] = \int_{-\infty}^{\infty} \boldsymbol{N}(x \mid \mu, \sigma^2)\, x^2\, dx = \mu^2 + \sigma^2$$

… (1.50)

[カルバック – ライブラーダイバージェンス]

$$\begin{aligned}\mathrm{KL}(p \| q) &= -\int p(\mathbf{x}) \ln q(\mathbf{x})\, d\mathbf{x} - \left(-\int p(\mathbf{x}) \ln p(\mathbf{x})\, d\mathbf{x}\right) \\ &= -\int p(\mathbf{x}) \ln\left\{\frac{q(\mathbf{x})}{p(\mathbf{x})}\right\} d\mathbf{x}\end{aligned}$$

… (1.113)

[解]

式 (1.113) より、2 つのガウス分布 $p(x)$, $q(x)$ 間のカルバック – ライブラーダイバージェンスは、以下のように表される。

$$\begin{aligned}\mathrm{KL}(p \| q) &= -\int p(x) \ln\left\{\frac{q(x)}{p(x)}\right\} dx \\ &= -\int N(x \mid \mu, \sigma^2) \ln\left\{\frac{N(x \mid m, s^2)}{N(x \mid \mu, \sigma^2)}\right\} dx\end{aligned}$$

… ※

ここで、上記の式の対数部分について整理する。ガウス分布は、式 (1.46) で与えられるので、

$$\ln\left\{\frac{N(\,x\mid m,\ s^2\,)}{N(\,x\mid \mu,\ \sigma^2\,)}\right\}$$

$$= \ln\frac{1}{(\,2\pi s^2\,)^{1/2}}\exp\left\{-\frac{1}{2s^2}(\,x-m\,)^2\right\} - \ln\frac{1}{(\,2\pi\sigma^2\,)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2}(\,x-\mu\,)^2\right\}$$

$$= \left\{-\frac{1}{2}\ln 2\pi s^2 - \frac{(\,x-m\,)^2}{2s^2}\right\} - \left\{-\frac{1}{2}\ln 2\pi\sigma^2 - \frac{(\,x-\mu\,)^2}{2\sigma^2}\right\}$$

$$= \frac{1}{2}(\,\ln 2\pi\sigma^2 - \ln 2\pi s^2\,) - \frac{(\,x-m\,)^2}{2s^2} + \frac{(\,x-\mu\,)^2}{2\sigma^2}$$

$$= \frac{1}{2}\ln\frac{\sigma^2}{s^2} - \frac{(\,x-m\,)^2}{2s^2} + \frac{(\,x-\mu\,)^2}{2\sigma^2}$$

$$= \ln\frac{\sigma}{s} - \frac{(\,x-m\,)^2}{2s^2} + \frac{(\,x-\mu\,)^2}{2\sigma^2}$$

… ※※

のように整理できる。式 (1.113) で与えられるカルバック − ライブラーダイバージェンス KL ($p \parallel q$) は、確率分布 $p\,(\,x\,)$ に基づく期待値としてみなすことができるので、その期待値を $\mathrm{E}_p[\,\cdot\,]$ とすると、式 (1.34) より、式 (1.113) は期待値 $\mathrm{E}_p[\,\cdot\,]$ を用いて、

$$\mathrm{KL}\,(\,p \parallel q\,) = -\int p\,(\,x\,)\ln\left\{\frac{q\,(\,x\,)}{p\,(\,x\,)}\right\}dx$$

$$= -\mathrm{E}_p\left[\ln\frac{q\,(\,x\,)}{p\,(\,x\,)}\right]$$

と書き直すことができ、さらに式 ※※ より、

$$= -\mathrm{E}_p\left[\ln\frac{\sigma}{s} - \frac{(\,x-m\,)^2}{2s^2} + \frac{(\,x-\mu\,)^2}{2\sigma^2}\right]$$

$$= -\mathrm{E}_p\left[\ln\frac{\sigma}{s}\right] + \mathrm{E}_p\left[\frac{(\,x-m\,)^2}{2s^2}\right] - \mathrm{E}_p\left[\frac{(\,x-\mu\,)^2}{2\sigma^2}\right]$$

… ※※※

と書き直せる。上記の式の第一項については、x に関する項を含まないので、

$$\mathrm{E}_p\left[\ln\frac{\sigma}{s}\right] = \ln\frac{\sigma}{s}$$

となり、第二項については、

$$\mathrm{E}_p\left[\frac{(\,x-m\,)^2}{2s^2}\right] = \frac{\mathrm{E}_p[\,x^2\,] - 2m\,\mathrm{E}_p[\,x\,] + m^2}{2s^2}$$

より、ここでガウス分布の 1 次と 2 次のモーメントは、式 (1.49), (1.50) で与えられるから、上記の式は、

$$= \frac{(\,\mu^2+\sigma^2\,) - 2m\cdot\mu + m^2}{2s^2} = \frac{\sigma^2 + (\,\mu-m\,)^2}{2s^2}$$

となり、第三項についても同じようにして、

$$\mathrm{E}_p\left[\ \frac{(x-\mu)^2}{2\sigma^2}\ \right] = \frac{\mathrm{E}_p[\ x^2\ ] - 2\mu\,\mathrm{E}_p[\ x\ ] + \mu^2}{2\sigma^2}$$

より、ここでガウス分布の 1 次と 2 次のモーメントは、式（ 1.49 ），（ 1.50 ）で与えられるから、上記の式は、

$$= \frac{(\ \mu^2 + \sigma^2\ ) - 2\mu\cdot\mu + \mu^2}{2\sigma^2} = \frac{1}{2}$$

となるので、最終的に式 ※※※ は、

$$= -\ln\frac{\sigma}{s} + \frac{\sigma^2 + (\mu - m)^2}{2s^2} - \frac{1}{2}$$

$$= \ln\frac{s}{\sigma} + \frac{\sigma^2 + (\mu - m)^2}{2s^2} - \frac{1}{2}$$

となる。以上より、2 つのガウス分布 $p(x) = N(\ x\,|\,\mu,\ \sigma^2\ )$ と $q(x) = N(\ x\,|\,\mu,\ s^2\ )$ の間のカルバック − ライブラーダイバージェンス $\mathrm{KL}(\,p\,\|\,q\,)$ は、

$$\mathrm{KL}(\,p\,\|\,q\,) = \ln\frac{s}{\sigma} + \frac{\sigma^2 + (\mu - m)^2}{2s^2} - \frac{1}{2}$$

となることがわかる。