演習問題 2.16

2 つの確率変数 x_1 と x_2 を考える。これらは、それぞれ平均が μ_1 と μ_2で、精度が τ_1 と τ_2 のガウス分布に従うとする。このとき、変数 $x = x_1 + x_2$ の微分エントロピーの式を導出せよ。これには、まず、次の関係を用い、x の分布を求め、指数部分を平方完成する。

$$p(x) = \int_{-\infty}^{\infty} p(x \mid x_2)\, p(x_2)\, dx_2$$

… (2.284)

次に、これが 2 つのガウス分布のたたみ込みになっており、また、これ自体もガウス分布になっていることに注目する。最後に、1 変数のガウス分布のエントロピーの結果 (1.110) を利用する。

[微分エントロピー]

$$\mathrm{H}[\mathbf{x}] = -\int p(\mathbf{x}) \ln p(\mathbf{x})\, d\mathbf{x}$$

… (1.104)

[ガウス分布の微分エントロピー]

$$\mathrm{H}[x] = \frac{1}{2}\{1 + \ln(2\pi\sigma^2)\}$$

… (1.110)

エントロピー $\mathrm{H}[x]$ は、σ^2 の値が増え、分布 $p(x)$ が幅広くなるに連れて大きくなることがわかる。この結果はまた、微分エントロピーが離散確率変数の場合とは異なり、負になり得ることを示している。$\sigma^2 < 1/(2\pi e)$ のとき、式 (1.110) から、$\mathbf{H}(x) < 0$ となるからである。

[ガウス分布]

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

… (2.42)

[多変量ガウス分布]

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right\}$$

… (2.43)

ただし、$\boldsymbol{\mu}$ は D 次元ベクトル、$\boldsymbol{\Sigma}$ は $D \times D$ の共分散行列、$|\boldsymbol{\Sigma}|$ は $\boldsymbol{\Sigma}$ の行列式を表す。

[解]

2 つの確率変数 x_1 と x_2 を考え、これらが、それぞれ平均が μ_1 と μ_2で、精度が τ_1 と τ_2 のガウス分布に従うとき、変数 $x = x_1 + x_2$ の微分エントロピーの式を導出する。2 つの確率変数におけるガウス分布は、それぞれ

$$\begin{cases} p(x_1) = \boldsymbol{N}(x_1 \mid \mu_1,\ \tau_1^{-1}) = \dfrac{1}{(2\pi\tau_1^{-1})^{1/2}} \exp\left\{ -\dfrac{1}{2\tau_1^{-1}}(x_1 - \mu_1)^2 \right\} \\ p(x_2) = \boldsymbol{N}(x_2 \mid \mu_2,\ \tau_2^{-1}) = \dfrac{1}{(2\pi\tau_2^{-1})^{1/2}} \exp\left\{ -\dfrac{1}{2\tau_2^{-1}}(x_2 - \mu_2)^2 \right\} \end{cases}$$

… ※

で与えられる。式 (2.284) より、$p(x)$ は、

$$p(x) = \int_{-\infty}^{\infty} p(x \mid x_2)\, p(x_2)\, dx_2$$

であり、いま、x は $x = x_1 + x_2$ で与えられるので、式 ※ で表される $p(x_1)$ より、

$$p(x_1) = \frac{1}{(2\pi\tau_1^{-1})^{1/2}} \exp\left\{ -\frac{1}{2\tau_1^{-1}}(x - \mu_1 - x_2)^2 \right\} = p(x \mid x_2)$$

と表されるので、式 (2.284) は、

$$p(x) = \int_{-\infty}^{\infty} p(x \mid x_2)\, p(x_2)\, dx_2$$

$$= \int_{-\infty}^{\infty} \frac{1}{(2\pi\tau_1^{-1})^{1/2}} \exp\left\{ -\frac{1}{2\tau_1^{-1}}(x - \mu_1 - x_2)^2 \right\} \cdot \frac{1}{(2\pi\tau_2^{-1})^{1/2}} \exp\left\{ -\frac{1}{2\tau_2^{-1}}(x_2 - \mu_2)^2 \right\} dx_2$$

$$= \left(\frac{\tau_1}{2\pi}\right)^{1/2} \left(\frac{\tau_2}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} \exp\left\{ -\frac{\tau_1}{2}(x - \mu_1 - x_2)^2 - \frac{\tau_2}{2}\left(x_2 - \mu_2\right)^2 \right\} dx_2$$

… ※※

と書き直せる。

式 (1.110) より、ガウス分布のエントロピーは分散、あるいは精度によって決定し、平均とは独立しており、さらに最終解は、$p(x)$ のガウス分布になることから、その精度のみを評価するだけで良いことがわかる。このことは、正規化係数などを無視することで、計算を簡潔化できることを示している。

ここで、上記の式を x_2 について平方完成するために、上記の式の指数部分を整理する。

$$-\frac{\tau_1}{2}(x - \mu_1 - x_2)^2 - \frac{\tau_2}{2}(x_2 - \mu_2)^2$$

$$-\frac{\tau_1}{2} x_2^2 - \frac{\tau_1}{2} \cdot 2(x - \mu_1) x_2 - \frac{\tau_1}{2}(x - \mu_1)^2 - \frac{\tau_2}{2} x_2^2 - \frac{\tau_2}{2} \cdot 2\mu_2 x_2 - \frac{\tau_2}{2}\mu_2^2$$

$$-\frac{1}{2}(\tau_1 + \tau_2) x_2^2 + \{\tau_1(x - \mu_1) + \tau_2\mu_2\} x_2$$

これを平方完成すると、

$$-\frac{1}{2}(\tau_1 + \tau_2) \left\{ x_2^2 - 2 \cdot \frac{\tau_1(x - \mu_1) + \tau_2\mu_2}{\tau_1 + \tau_2} \cdot x_2 \right\}$$

$$= -\frac{1}{2}(\tau_1 + \tau_2) \left[\left\{ x_2 - \frac{\tau_1(x - \mu_1) + \tau_2\mu_2}{\tau_1 + \tau_2} \right\}^2 - \left\{ \frac{\tau_1(x - \mu_1) + \tau_2\mu_2}{\tau_1 + \tau_2} \right\}^2 \right]$$

$$
= -\frac{1}{2}(\tau_1 + \tau_2)\left\{ x_2 - \frac{\tau_1(x - \mu_1) + \tau_2\mu_2}{\tau_1 + \tau_2} \right\}^2 + \frac{\{\tau_1(x - \mu_1) + \tau_2\mu_2\}^2}{2(\tau_1 + \tau_2)}
$$

となる。このことから、式 ※※ で表される $p(x)$ は、

$$
= \left(\frac{\tau_1}{2\pi}\right)^{1/2}\left(\frac{\tau_2}{2\pi}\right)^{1/2}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}(\tau_1 + \tau_2)\left\{ x_2 - \frac{\tau_1(x - \mu_1) + \tau_2\mu_2}{\tau_1 + \tau_2} \right\}^2 + \frac{\{\tau_1(x - \mu_1) + \tau_2\mu_2\}^2}{2(\tau_1 + \tau_2)} \right] dx_2
$$

… ※※※

と書き表せる。これが 2 つのガウス分布のたたみ込みになっており、また、これ自体もガウス分布になっていることに注目する。上記の式の第一項は、単純に独立変数 x の式となっており、第二項指数部は展開したときに、x^2 を伴う。

[式 ※※※ の第二項指数部の展開 ⇒ x^2 項の取得]

$$
\frac{\{\tau_1(x - \mu_1) + \tau_2\mu_2\}^2}{2(\tau_1 + \tau_2)} = \frac{\{\tau_1(x - \mu_1)\}^2 + 2\tau_1\tau_2\mu_2(x - \mu_1) + \tau_2^2\mu_2^2}{2(\tau_1 + \tau_2)}
$$

$$
= \frac{\tau_1^2x^2 - 2\tau_1\mu_1x + \tau_1\mu_1^2 + 2\tau_1\tau_2\mu_2(x - \mu_1) + \tau_2^2\mu_2^2}{2(\tau_1 + \tau_2)}
$$

$$
= \frac{\tau_1^2x^2 + 2\tau_1(\tau_2\mu_2 - \mu_1)x + \tau_1\mu_1^2 + 2\tau_1\tau_2\mu_1\mu_2 + \tau_2^2\mu_2^2}{2(\tau_1 + \tau_2)}
$$

$$
= \frac{\tau_1^2}{2(\tau_1 + \tau_2)}x^2 + \frac{\tau_1(\tau_2\mu_2 - \mu_1)}{\tau_1 + \tau_2}x + \frac{\tau_1\mu_1^2 + 2\tau_1\tau_2\mu_1\mu_2 + \tau_2^2\mu_2^2}{2(\tau_1 + \tau_2)}
$$

x の精度は、指数部分中の x^2 の係数項で直接与えられるので、この部分だけを考慮する。同様に、式 ※※ は、x について平方完成したものと見なせる。式 ※※ からも、x^2 の項が得られる。これら 2 つの x^2 項を合わせると、

$$
\underbrace{-\frac{\tau_1}{2}x^2}_{\text{式 ※※ より}} + \underbrace{\frac{\tau_1^2}{2(\tau_1 + \tau_2)}x^2}_{\text{式 ※※※ の第二項より}} = -\frac{1}{2}\frac{\tau_1\tau_2}{\tau_1 + \tau_2}x^2
$$

となる。これから、確率変数 x は、精度 $\tau_1\tau_2/(\tau_1 + \tau_2)$ をもつことがわかる。

この精度解は、2 つの線形ガウス分布のたたみ込みの一般解からも得ることができる。

よって、式 (1.110) より、$\sigma^2 = (\tau_1\tau_2/(\tau_1 + \tau_2))^{-1}$ であるから、変数 $x = x_1 + x_2$ の微分エントロピーは、

$$
\mathrm{H}[x] = \frac{1}{2}\ln\left\{ \frac{2\pi(\tau_1 + \tau_2)}{\tau_1\tau_2} \right\}
$$

で与えられることがわかる。