PRML 演習 1.20-1.21

@amoO_O

（演習 1.20）
以下で、x は D 次元ベクトルである。
演習問題(1.18)の式(1.142)：

$$\prod_{i=1}^{D}\int_{-\infty}^{\infty}e^{-x_i^2}dx_i = S_D\int_0^{\infty}e^{-r^2}r^{D-1}dr$$

は、原点中心、半径 R の D 次元球内に限定しても等式が成り立ち、以下のように書ける：

$$\int_{\|x\|\leq R}f(\|x\|)dx = S_D\int_0^R f(r)r^{D-1}dr$$

「半径 r にある厚さ ε の薄皮に関して x の確率密度の積分をとる」とは、「半径 r〜半径 r+ε」までの積分と解釈できる。したがって、

$$\begin{aligned}
\int_{r\leq\|x\|\leq r+\epsilon}\frac{1}{(2\pi\sigma^2)^{D/2}}exp\Big(-\frac{\|x\|^2}{2\sigma^2}\Big)dx^{to} &= \Big(\int_{\|x\|\leq r+\epsilon}-\int_{\|x\|\leq r}\Big)\frac{1}{(2\pi\sigma^2)^{D/2}}exp\Big(-\frac{\|x\|^2}{2\sigma^2}\Big)dx \\
&= S_D\Big(\int_0^{r+\epsilon}-\int_0^{r}\Big)\frac{1}{(2\pi\sigma^2)^{D/2}}exp\Big(-\frac{s^2}{2\sigma^2}\Big)s^{D-1}ds \\
&= \int_r^{r+\epsilon}\frac{S_D s^{D-1}}{(2\pi\sigma^2)^{D/2}}exp\Big(-\frac{s^2}{2\sigma^2}\Big)ds \\
&= \int_r^{r+\epsilon}p(s)ds \\
&\simeq p(r)\epsilon(\epsilon\ll 1)
\end{aligned}$$

次に p(r)を r について微分すると、

$$\begin{aligned}
&to\\
\frac{dp}{dr} &= \frac{S_D}{(2\pi\sigma^2)^{D/2}}\frac{d}{dr}(r^{D-1}exp(-\frac{r^2}{2\sigma^2}))\\
&= \frac{S_D}{(2\pi\sigma^2)^{D/2}}((D-1)r^{D-2}+r^{D-1}(-\frac{r}{\sigma^2}))exp(-\frac{r^2}{2\sigma^2})\\
&= \frac{S_D}{(2\pi\sigma^2)^{D/2}}((D-1)r^{D-2}-\frac{r^D}{\sigma^2})exp(-\frac{r^2}{2\sigma^2})
\end{aligned}$$

r>0 の時、定数項と exp、r^D-2 は 0 にならないので、

$$\begin{aligned}
&to\\
\left.\frac{dp}{dr}\right|_{r=\hat{r}} &= 0\\
\Leftrightarrow (D-1)-\frac{\hat{r}^2}{\sigma^2} &= 0\\
\Leftrightarrow \hat{r} &= \sqrt{D-1}\sigma\\
&\simeq \sqrt{D}\sigma
\end{aligned}$$

$$\begin{aligned}
&to\\
\frac{p(\hat{r}+\epsilon)}{p(\hat{r})} &= \frac{\frac{S_D}{(2\pi\sigma^2)^{D/2}}(\hat{r}+\epsilon)^{D-1}exp\left(-\frac{(\hat{r}+\epsilon)^2}{2\sigma^2}\right)}{\frac{S_D}{(2\pi\sigma^2)^{D/2}}\hat{r}^{D-1}exp\left(-\frac{\hat{r}^2}{2\sigma^2}\right)}\\
&= \frac{(\hat{r}+\epsilon)^{D-1}}{\hat{r}^{D-1}}exp\left(-\frac{(\hat{r}+\epsilon)^2}{2\sigma^2}+\frac{\hat{r}^2}{2\sigma^2}\right)\\
&= \left(1+\frac{\epsilon}{\hat{r}}\right)^{D-1}exp\left(-\frac{2\hat{r}\epsilon+\epsilon^2}{2\sigma^2}\right)
\end{aligned}$$

ここで右辺の(1+ε/r)^D-1 を先に「ほぼ 1」と判断してしまうと、exp の中が正しく評価できず、(1.149)を導けない。両辺 ln を取ると、

$$ln\left(p(\hat{r}+\epsilon)\right)-ln\left(p(\hat{r})\right)=(D-1)ln\left(1+\frac{\epsilon}{\hat{r}}\right)-\frac{\hat{r}\epsilon}{\sigma^2}-\frac{\epsilon^2}{2\sigma^2}$$

$$\begin{aligned}
&to\\
1 &\ll D\\
\epsilon &\ll \hat{r}\simeq\sqrt{D}\sigma\\
ln(1+x) &\simeq x-\frac{x^2}{2}+O(x^3)(x\ll 1)
\end{aligned}$$

に注意して上式を変形すると、

$$\begin{array}{rcl} to & & \\ ln\left(p(\hat{r}+\epsilon)\right)-ln\left(p(\hat{r})\right) & \simeq & (D-1)\left(\frac{\epsilon}{\hat{r}}-\frac{\epsilon^2}{2\hat{r}^2}\right)-\frac{\hat{r}\epsilon}{\sigma^2}-\frac{\epsilon^2}{2\sigma^2} \\ & \simeq & D\left(\frac{\epsilon}{\sqrt{D}\sigma}-\frac{\epsilon^2}{2D\sigma^2}\right)-\frac{\sqrt{D}\sigma\epsilon}{\sigma^2}-\frac{\epsilon^2}{2\sigma^2} \\ & \simeq & \frac{\sqrt{D}\epsilon}{\sigma}-\frac{\epsilon^2}{2\sigma^2}-\frac{\sqrt{D}\epsilon}{\sigma}-\frac{\epsilon^2}{2\sigma^2} \\ & = & -\frac{\epsilon^2}{\sigma^2} \end{array}$$

最後に、

$$\|\hat{x}\|=\hat{r}$$

なる x^ を使って、

$$\begin{array}{rcl} to & & \\ \frac{p(0)}{p(x)|_{x=\hat{x}}} & = & \frac{\frac{1}{(2\pi\sigma^2)^{D/2}}exp(0)}{\frac{1}{(2\pi\sigma^2)^{D/2}}exp\left(-\frac{\|\hat{x}\|^2}{2\sigma^2}\right)} \\ & = & exp\left(\frac{\|\hat{x}\|^2}{2\sigma^2}\right) \\ & = & exp\left(\frac{\hat{r}^2}{2\sigma^2}\right) \\ & \simeq & exp\left(\frac{D\sigma^2}{2\sigma^2}\right) \\ & = & exp\left(\frac{D}{2}\right) \end{array}$$

（演習 1.21）

$$
\begin{aligned}
&to \\
&\quad 0 \leq a \leq b \\
&\Rightarrow \quad a^{1/2} \leq b^{1/2} \\
&\Rightarrow \quad a \leq (ab)^{1/2}
\end{aligned}
$$

p.39 の図 1.24 より、クラス分類問題の決定領域を誤識別率が最小になるように選んだ場合、下記が成り立つ。

$$
\begin{aligned}
x \in R_1 \Rightarrow p(x, C_2) \leq p(x, C_1) \\
x \in R_2 \Rightarrow p(x, C_1) \leq p(x, C_2)
\end{aligned}
$$

これに先の結果を適用して、

$$
\begin{aligned}
x \in R_1 \Rightarrow p(x, C_2) \leq \{p(x, C_1)p(x, C_2)\}^{1/2} \\
x \in R_2 \Rightarrow p(x, C_1) \leq \{p(x, C_1)p(x, C_2)\}^{1/2}
\end{aligned}
$$

それぞれの領域で積分し、加えると、

$$
\begin{aligned}
&to \\
p(Ayamari) &= \int_{R_1} p(x, C_2)dx + \int_{R_2} p(x, C_1)dx \\
&\leq \int_{R_1} \{p(x, C_1)p(x, C_2)\}^{1/2} dx + \int_{R_2} \{p(x, C_1)p(x, C_2)\}^{1/2} dx \\
&= \int \{p(x, C_1)p(x, C_2)\}^{1/2} dx
\end{aligned}
$$

※日本語が打てない