演習問題 1.16

演習問題 1.15 で、M 次の D 次元多項式の独立なパラメータの数が、式 (1.35) となることを証明した。次に、M 次までのすべての項における独立パラメータの総数 $N(D,M)$ を求めよう。まず、$N(D,M)$ が

$$N(D,M)=\sum_{m=0}^{M} n(D,m) \qquad \cdots (1.38)$$

を満たすことを示せ。ただし、$N(D,M)$ は、m 次の項における独立パラメータの数である。式 (1.37) の結果と、数学的帰納法により、

$$N(D,M)=\frac{(D+M)!}{D!\,M!} \qquad \cdots (1.39)$$

を示せ。これを示すには、まず、$M=0$ と任意の $D \gg 1$ について成り立つことを証明してから、M 次で成り立つなら、$M+1$ 次で成り立つことを示せばよい。最後にスターリングの近似式、つまり n が大きいとき、

$$n! \cong n^n e^{-n} \qquad \cdots (1.40)$$

が成り立つことを用いて、$D \gg M$ のとき、$N(D,M)$ が D^M で大きくなり、$M \gg D$ のときは、M^D で大きくなることを示せ。D 次元の 3 次多項式 ($M=3$) を考え、独立パラメータの総数を (ⅰ) $D=10$ (ⅱ) $D=100$ のそれぞれの場合について数値的に評価せよ。これらは典型的なスケールおよび中スケールの機械学習の応用に対応する。

[M 次の D 次元多項式の独立なパラメータの数]

$$n(D,M)=\sum_{i=1}^{D} n(i,M-1) \qquad \cdots (1.35)$$

$$n(D,M)=\frac{(D+M-1)!}{(D-1)!\,M!} \qquad \cdots (1.37)$$

[解]

まず、$N(D,M)$ が式 (1.38) を満たすことを示す。M 次における D 次元多項式の独立なパラメータの数が $n(D,M)$ で表されるから、M 次までのすべての項における独立なパラメータの総数 $N(D,M)$ は、

$$N(D,M)=n(D,0)+n(D,2)+\cdots+n(D,M)=\sum_{m=0}^{M} n(D,m)$$

と表せる。

次に、式 (1.37) の結果と、数学的帰納法により、式 (1.39) を証明する。ここでは、まず、$M=0$ と任意の $D \gg 1$ について成り立つことを証明してから、M 次で成り立つなら、$M+1$ 次で成り立つことを示す。

（i） $M=0$ と任意の $D \gg 1$ のとき、式 (1.39) より、

$$\text{式 (1.39) の左辺} = N(D,\ 0) = \frac{(D+0-1)!}{(D-1)!\,0!} = 1$$

$$\text{式 (1.39) の右辺} = \sum_{m=0}^{0} n(D,\ 0) = \frac{(D+0)!}{D!\,0!} = 1$$

となるため、式 (1.39) は成り立つ。

（ii） $M=M$ と任意の $D \gg 1$ の場合に、式 (1.39) が成立すると仮定すると、$M=M+1$ と任意の $D \gg 1$ のとき、式 (1.39) の左辺は、

$$\text{式 (1.39) の左辺} = N(D,\ M+1)$$

であり、式 (1.38) より、

$$= \sum_{m=0}^{M+1} n(D,\ m)$$

$$= n(D,\ M+1) + \sum_{m=0}^{M} n(D,\ m)$$

$$= n(D,\ M+1) + N(D,\ M)$$

と分解できる。ここで、上記の式に、式 (1.37), (1.39) を代入すると、

$$= \frac{(D+M)!}{(D-1)!\,(M+1)!} + \frac{(D+M)!}{D!\,M!}$$

$$= \frac{D(D+M)! + (M+1)(D+M)!}{D!\,(M+1)!}$$

$$= \frac{(D+M+1)(D+M)!}{D!\,(M+1)!}$$

$$= \frac{\{(D+M)+1\}(D+M)!}{D!\,(M+1)!}$$

$$= \frac{(D+M+1)!}{D!\,(M+1)!}$$

$$= \frac{(D+M+1)!}{D!\,(M+1)!}$$

と整理できる。$M=M+1$ と任意の $D \gg 1$ のとき、式 (1.39) の右辺は、

$$\text{式 (1.39) の右辺} = \frac{(D+M+1)!}{D!\,(M+1)!}$$

であるから、式 (1.39) は成り立つ。

以上（i）,（ii）より、数学的帰納法を用いて、$M \gg 0$ と任意の $D \gg 1$ について、式（1.39）が成り立つことが証明された。

最後に、n が大きいとき、スターリングの近似式（1.40）が成り立つことを用いて、$D \gg M$ のとき、$N(D, M)$ が D^M で大きくなり、$M \gg D$ のときは、$N(D, M)$ が M^D で大きくなることを示す。

（i）$D \gg M$ のとき、スターリングの近似式（1.40）において、$n = D$ として、

$$D! \cong D^D e^{-D}$$

となるから、以下の式が成り立つ。

$$(D+M)! \cong (D+M)^{(D+M)} e^{-(D+M)}$$

ここで、今、D が M よりも十分に大きいとしているので、上記の式はさらに、

$$\cong D^{(D+M)} e^{-D}$$

$$= D^D D^M e^{-D}$$

と書き直せる。よって、式（1.39）は、

$$N(D, M) = \frac{(D+M)!}{D!\,M!}$$

$$= \frac{D^D D^M e^{-D}}{D^D e^{-D} M!}$$

$$= \frac{D^M}{M!}$$

となるので、$D \gg M$ の場合、$N(D, M)$ が D^M で大きくなることが示せた。

（ii）$M \gg D$ のとき

上記と同様、スターリングの近似式（1.40）において、$n = M$ として、式（1.39）を整理することにより、$M \gg D$ の場合、$N(D, M)$ が M^D で大きくなることが示せる。

D 次元の 3 次多項式（$M = 3$）を考え、独立パラメータの総数を（i）$D = 10$（ii）$D = 100$ のそれぞれの場合について数値的に評価しておく。スターリングの近似式と、それを用いて式（1.39）で表される $N(D, M)$ を整理して求めた独立パラメータの総数を以下に示す。

M, D の値	スターリングの近似式	真値
$M = 3$, $D = 10$	$10^3 / 3! \cong 167$	$N(10, 3) = 286$
$M = 3$, $D = 100$	$100^3 / 3! \cong 166667$	$N(100, 3) = 176851$