演習問題 2.09

この演習問題では、ディクリレ分布が正規化されていることを、帰納法によって証明する。ディクリレ分布で $M = 2$ とした特殊な場合であるベータ分布が正規化されていることは、演習問題 2.5 ですでに示した。次に、ディクリレ分布が $M - 1$ 変数の場合に正規化されていることの仮定の下、M 変数でも正規化されていることを証明する。これにはまず、M 変数のディクリレ分布から、$\sum_{k=1}^{M} \mu_k = 1$ の制約を用いて、μ_M を除去し、ディクリレ分布を

$$p_M(\mu_1, \cdots, \mu_{M-1}) = C_M \prod_{k=1}^{M-1} \mu_k^{\alpha_k - 1} \left(1 - \sum_{j=1}^{M-1} \mu_j \right)^{\alpha_M - 1}$$

… (2.272)

と書く。すると、あとは C_M を表す式を求めればよい。これには、この式を、範囲に注意しつつ μ_{M-1} について積分し、さらに、この積分の範囲が 0 から 1 になるように変数を置換する。C_{M-1} での結果が正しいと仮定し、式 (2.265) を用いて、C_M の式を導出せよ。

[ディクリレ分布]

○ 共役性をもつ多項分布

$$\mathrm{Dir}(\boldsymbol{\mu} \mid \boldsymbol{\alpha}) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_M)} \prod_{k=1}^{M} \mu_k^{\alpha_k - 1}$$

$$\left(= \underbrace{\frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_M)}{\Gamma(\alpha_1)\Gamma(\alpha_2) \cdots \Gamma(\alpha_M)}}_{\text{正規化係数}} \prod_{k=1}^{M} \mu_k^{\alpha_k - 1} \right)$$

… (2.38)

ただし、$0 \le \mu_k \le 1$ かつ $\sum_k \mu_k = 1$ である。ここで、$\alpha_1, \cdots, \alpha_M$ は、この分布のパラメータで、$\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_M)^{\mathrm{T}}$ を表す。また、α_0 は、

$$\alpha_0 = \sum_{k=1}^{K} \alpha_k$$

… (2.39)

である。

[ベータ分布の正規化条件]

$$\int_0^1 \mu^{a-1} (1 - \mu)^{b-1} d\mu = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}$$

… (2.265)

[式 (2.272) の証明]

式 (2.38) より、ここでのディリクレ分布 $p_M(\mu_1, \cdots, \mu_M)$ は、

$$p_M(\mu_1, \cdots, \mu_M) = C_M \prod_{k=1}^{M} \mu_k^{\alpha_k - 1}$$

と書き表すことができる。ただし、C_M は正規化係数を表す。ここで、μ_M を直積の外に出

すと、上記の式は、

$$p_M(\mu_1, \cdots, \mu_{M-1}) = C_M \prod_{k=1}^{M-1} \mu_k^{\alpha_k - 1} \cdot \mu_k^{\alpha_M - 1}$$

と書き直せる。ここで、制約条件 $\sum_j^M \mu_j = 1$ より、$\mu_M = 1 - \sum_j^{M-1} \mu_j$ が成り立つので、上記の式は、

$$= C_M \prod_{k=1}^{M-1} \mu_k^{\alpha_k - 1} \left(1 - \sum_{j=1}^{M-1} \mu_j \right)^{\alpha_M - 1}$$

となる。よって、式（2.272）が導かれた。

［解］

ディクリレ分布が $M-1$ 変数の場合に正規化されていることの仮定の下、数学的帰納法を用いて、M 変数でも正規化されていることを証明する。

(i) $M=2$ のとき、ディクリレ分布（2.38）は、$\sum_{k=1}^M \mu_k = 1$ より、$\mu_1 = \mu$, $\mu_2 = 1-\mu$ が成り立つことと、$\alpha_1 = a$, $\alpha_2 = b$ とおくことにより、ベータ分布と同じ形になるので、演習問題 2.5 より、ディクリレ分布（2.37）は正規化されていることが示されている。

(ii) $M=M-1$ のとき、ディクリレ分布（2.38）が正規化されていると仮定した下で、$M=M$ のときもディクリレ分布（2.38）が正規化されていることを証明する。題意より、式（2.272）を積分区間に注意しつつ μ_{M-1} について積分する。

$$(\quad p_{M-1}(\mu_1, \cdots, \mu_{M-2}) = \quad)$$

$$\int_0^{\mu_{M-1}} p_M(\mu_1, \cdots, \mu_{M-1})\, d\mu_{M-1}$$

すると、$M=M-1$ のとき、制約条件 $\sum_{j=1}^{M-1} \mu_j = 1$ より、$\mu_{M-1} = 1 - \sum_{j=1}^{M-2} \mu_j$ が成り立つため、

$$= \int_0^{1-\sum_{j=1}^{M-2}\mu_j} p_M(\mu_1, \cdots, \mu_{M-1})\, d\mu_{M-1}$$

$$= \int_0^{1-\sum_{j=1}^{M-2}\mu_j} C_M \prod_{k=1}^{M-1} \mu_k^{\alpha_k - 1} \left(1 - \sum_{j=1}^{M-1} \mu_j \right)^{\alpha_M - 1} d\mu_{M-1}$$

… ※

と書き表せる。さらに、この積分の範囲が 0 から 1 になるように変数変換をする。ここで、$d\mu_{M-1} = dt$ とおくと、μ_{M-1} と t の積分区間の対応関係は、

t の積分区間	0	…	1
μ_{M-1} の積分区間	0	…	$1 - \sum_{j=1}^{M-2} \mu_j$

となるため、

$$t : \mu_{M-1} = 1 : 1 - \sum_{j=1}^{M-2} \mu_j \Leftrightarrow \mu_{M-1} = t\left(1 - \sum_{j=1}^{M-2} \mu_j\right)$$

… ※※

より、μ_{M-1} と t 間に、上記の式のような線形変換の関係があることがわかる。この変換を用いると、式 ※ は、

$$= \int_0^1 C_M \underline{\prod_{k=1}^{M-1} \mu_k^{\alpha_k - 1}} \left(1 - \sum_{j=1}^{M-1} \mu_j\right)^{\alpha_M - 1} dt$$

… ※※※

と書き直せ、さらに上記中の $\prod_{k=1}^{M-1} \mu_k^{\alpha_k - 1}$ は、

$$\prod_{k=1}^{M-1} \mu_k^{\alpha_k - 1} = \prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \cdot \mu_{M-1}^{\alpha_{M-1} - 1}$$

と展開でき、ここで、変数変換式 ※※ より、

$$= \prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \left\{ t\left(1 - \sum_{j=1}^{M-2} \mu_j\right)\right\}^{\alpha_{M-1} - 1}$$

となることから、式 ※※※ は、

$$= \int_0^1 C_M \prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \left\{ t\left(1 - \sum_{j=1}^{M-2} \mu_j\right)\right\}^{\alpha_{M-1} - 1} \left(1 - \sum_{j=1}^{M-1} \mu_j\right)^{\alpha_M - 1} dt$$

$$= C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1}\right) \int_0^1 \left\{ t\left(1 - \sum_{j=1}^{M-2} \mu_j\right)\right\}^{\alpha_{M-1} - 1} \left(1 - \sum_{j=1}^{M-1} \mu_j\right)^{\alpha_M - 1} dt$$

$$= C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1}\right) \int_0^1 t^{\alpha_{M-1} - 1} \left(1 - \sum_{j=1}^{M-2} \mu_j\right)^{\alpha_{M-1} - 1} \left(1 - \sum_{j=1}^{M-1} \mu_j\right)^{\alpha_M - 1} dt$$

$$= C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1}\right) \left(1 - \sum_{j=1}^{M-2} \mu_j\right)^{\alpha_{M-1} - 1} \int_0^1 t^{\alpha_{M-1} - 1} \underline{\left(1 - \sum_{j=1}^{M-1} \mu_j\right)^{\alpha_M - 1}} dt$$

… ※※※※

と整理できる。また、上記中の $1 - \sum_{j=1}^{M-1} \mu_j$ は、

$$1 - \sum_{j=1}^{M-1} \mu_j = 1 - \left(\sum_{j=1}^{M-2} \mu_j + \mu_{M-1}\right)$$

と展開できるので、変数変換式 ※※ より、上記の式は、

$$= 1 - \left\{ \sum_{j=1}^{M-2} \mu_j + t \left(1 - \sum_{j=1}^{M-2} \mu_j \right) \right\}$$

$$= 1 - \sum_{j=1}^{M-2} \mu_j - t + t \sum_{j=1}^{M-2} \mu_j$$

$$= (1 - t) - (1 - t) \sum_{j=1}^{M-2} \mu_j$$

$$= (1 - t) \left(1 - \sum_{j=1}^{M-2} \mu_j \right)$$

と整理できる。このことから、式 ※※※※ は、

$$= C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \right) \left(1 - \sum_{j=1}^{M-2} \mu_j \right)^{\alpha_{M-1} - 1} \int_0^1 t^{\alpha_{M-1} - 1} (1 - t)^{\alpha_M - 1} \left(1 - \sum_{j=1}^{M-2} \mu_j \right)^{\alpha_M - 1} dt$$

と書き直せ、これを整理すると、

$$= C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \right) \left(1 - \sum_{j=1}^{M-2} \mu_j \right)^{\alpha_{M-1} - 1} \left(1 - \sum_{j=1}^{M-2} \mu_j \right)^{\alpha_M - 1} \int_0^1 t^{\alpha_{M-1} - 1} (1 - t)^{\alpha_M - 1} dt$$

$$= C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \right) \left(1 - \sum_{j=1}^{M-2} \mu_j \right)^{\alpha_{M-1} + \alpha_M - 1} \int_0^1 t^{\alpha_{M-1} - 1} (1 - t)^{\alpha_M - 1} dt$$

となる。ここで、式 (2.265) が証明されているから、上記の式は、

$$= C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \right) \left(1 - \sum_{j=1}^{M-2} \mu_j \right)^{\alpha_{M-1} + \alpha_M - 1} \frac{\Gamma(\alpha_{M-1}) \Gamma(\alpha_M)}{\Gamma(\alpha_{M-1} + \alpha_M)}$$

… ※※※※※

と書き表せる。ここで、$M = M - 1$ のとき、ディクリレ分布 (2.38) が正規化されていると仮定しているので、

$$p_{M-1}(\mu_1, \cdots, \mu_{M-2})$$

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{M-1})}{\Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_{M-1})} \prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \left(1 - \sum_{j=1}^{M-2} \mu_j \right)^{\alpha_{M-1} - 1}$$

が保証される。また、上記の式において、$\alpha_{M-1} = \alpha_{M-1} + \alpha_M$ とおくと、

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{M-1} + \alpha_M)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_{M-1} + \alpha_M)} \prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \left(1 - \sum_{j=1}^{M-2} \mu_j\right)^{\alpha_{M-1} + \alpha_M - 1}$$

となり、この式が式 ※※※※※ と等しくなるので、

$$C_M \left(\prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1}\right)\left(1 - \sum_{j=1}^{M-2} \mu_j\right)^{\alpha_{M-1} + \alpha_M - 1} \frac{\Gamma(\alpha_{M-1})\Gamma(\alpha_M)}{\Gamma(\alpha_{M-1} + \alpha_M)}$$

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{M-1} + \alpha_M)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_{M-1} + \alpha_M)} \prod_{k=1}^{M-2} \mu_k^{\alpha_k - 1} \left(1 - \sum_{j=1}^{M-2} \mu_j\right)^{\alpha_{M-1} + \alpha_M - 1}$$

より、これを C_M について解くと、

$$C_M = \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{M-1} + \alpha_M)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_{M-1} + \alpha_M)} \cdot \frac{\Gamma(\alpha_{M-1} + \alpha_M)}{\Gamma(\alpha_{M-1})\Gamma(\alpha_M)}$$

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{M-1} + \alpha_M)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_{M-1})\Gamma(\alpha_M)}$$

となる。よって、ディクリレ分布が $M - 1$ 変数の場合に正規化されていることの仮定の下、M 変数でも正規化されていることが示せた。

以上（i）,（ii）より、数学的帰納法を用いて、ディクリレ分布が正規化されていることが証明された。