パターン認識と機械学習

2.4 指数型分布族

吉田拓夢

指数型分布族

$$p(\mathbf{x}|\boldsymbol{\eta}) = h(\mathbf{x})g(\boldsymbol{\eta})\exp\{\boldsymbol{\eta}^{\mathrm{T}}\mathbf{u}(\mathbf{x})\}$$

- **x** : ベクトル/スカラー、離散/連続　どれでも可
- **η**:自然パラメータ
- **u**(**x**):**x**の任意の関数
- g(**η**):分布を正規化するパラメータ(以下を満たす)

$$g(\boldsymbol{\eta}) \int h(\mathbf{x})\exp\{\boldsymbol{\eta}^{\mathrm{T}}\mathbf{u}(\mathbf{x})\}\mathrm{d}\mathbf{x} = 1$$

ベルヌーイ分布

$$p(x|\mu) = \mathrm{Bern}(x|\mu) = \mu^x (1-\mu)^{1-x}$$

↓右辺の対数をとって、さらにその指数をとる

$$\begin{aligned} p(x|\mu) &= \exp\{x \ln \mu + (1-x)\ln(1-\mu)\} \\ &= (1-\mu)\exp\left\{\ln\left(\frac{\mu}{1-\mu}\right)x\right\} \end{aligned}$$

ベルヌーイ分布

指数型分布族の標準形と比較して

$$\eta = \ln\left(\frac{\mu}{1-\mu}\right)$$

これをμについて解くとμ=σ(η)

↓σ(η)はロジスティックシグモイド関数

$$\sigma(\eta) = \frac{1}{1+\exp(-\eta)}$$

ベルヌーイ分布

$$p(\mathbf{x}|\boldsymbol{\eta}) = h(\mathbf{x})g(\boldsymbol{\eta})\exp\{\boldsymbol{\eta}^{\mathrm{T}}\mathbf{u}(\mathbf{x})\}$$

↓指数型分布族の標準形で表したベルヌーイ分布

$$\Rightarrow \quad p(x|\eta) = \sigma(-\eta)\exp(\eta x)$$

$$\begin{aligned} u(x) &= x \\ h(x) &= 1 \\ g(\eta) &= \sigma(-\eta) \end{aligned}$$

$$\eta = \ln\left(\frac{\mu}{1-\mu}\right)$$

$$\sigma(\eta) = \frac{1}{1+\exp(-\eta)}$$

多項分布-1

$$p(\mathbf{x}|\boldsymbol{\mu}) = \prod_{k=1}^{M} \mu_k^{x_k} = \exp\left\{\sum_{k=1}^{M} x_k \ln \mu_k\right\}$$

$$\Rightarrow \quad p(\mathbf{x}|\boldsymbol{\eta}) = \exp(\boldsymbol{\eta}^{\mathrm{T}}\mathbf{x})$$

$$\eta_k = \ln \mu_k, \boldsymbol{\eta} = (\eta_1, \ldots, \eta_M)^{\mathrm{T}}$$

$$\begin{aligned} \mathbf{u}(\mathbf{x}) &= \mathbf{x} \\ h(\mathbf{x}) &= 1 \\ g(\boldsymbol{\eta}) &= 1 \end{aligned}$$

$$\sum_{k=1}^{M} \mu_k = 1$$

制約あり
η_kは独立でない

多項分布-2

- パラメータμ_kの内どのM-1個の値が与えられても残りのパラメータが決定する
- M-1個のみで分布を表した方が便利な場合もある (これら残りのパラメータはまだ以下の制約がある)

$$0 \leq \mu_k \leq 1, \quad \sum_{k=1}^{M-1} \mu_k \leq 1$$

多項分布-2

$$\sum_{k=1}^{M} \mu_k = 1$$ 制約(←)を用いて式変形

$$\exp\left\{\sum_{k=1}^{M} x_k \ln \mu_k\right\}$$

$$= \exp\left\{\sum_{k=1}^{M-1} x_k \ln \mu_k + \left(1 - \sum_{k=1}^{M-1} x_k\right) \ln\left(1 - \sum_{k=1}^{M-1} \mu_k\right)\right\}$$

$$= \exp\left\{\sum_{k=1}^{M-1} x_k \ln\left(\frac{\mu_k}{1 - \sum_{j=1}^{M-1} \mu_j}\right) + \ln\left(1 - \sum_{k=1}^{M-1} \mu_k\right)\right\}$$

多項分布-2

対応付けると

$$\ln\left(\frac{\mu_k}{1-\sum_j \mu_j}\right) = \eta_k$$

μ_kについて解くと

$$\mu_k = \frac{\exp(\eta_k)}{1+\sum_j \exp(\eta_j)}$$

↑ソフトマックス関数/正規化指数関数

多項分布-2

$$p(\mathbf{x}|\boldsymbol{\eta}) = \left(1 + \sum_{k=1}^{M-1} \exp(\eta_k)\right)^{-1} \exp(\boldsymbol{\eta}^{\mathrm{T}}\mathbf{x})$$

$$\eta_k = \ln\left(\frac{\mu_k}{1 - \sum_j \mu_j}\right), \boldsymbol{\eta} = (\eta_1, \ldots, \eta_{M-1})^{\mathrm{T}}$$

$$\mathbf{u}(\mathbf{x}) = \mathbf{x}$$

$$h(\mathbf{x}) = 1$$

$$g(\boldsymbol{\eta}) = \left(1 + \sum_{k=1}^{M-1} \exp(\eta_k)\right)^{-1}$$

ガウス分布

$$
\begin{aligned}
p(x|\mu,\sigma^2) &= \frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} \\
&= \frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2}x^2+\frac{\mu}{\sigma^2}x-\frac{1}{2\sigma^2}\mu^2\right\}
\end{aligned}
$$

$$
\Rightarrow \quad p(\mathbf{x}|\boldsymbol{\eta}) = h(\mathbf{x})g(\boldsymbol{\eta})\exp\{\boldsymbol{\eta}^{\mathrm{T}}\mathbf{u}(\mathbf{x})\}
$$

$$
\begin{aligned}
\boldsymbol{\eta} &= \begin{pmatrix} \mu/\sigma^2 \\ -1/2\sigma^2 \end{pmatrix} \\
\mathbf{u}(x) &= \begin{pmatrix} x \\ x^2 \end{pmatrix}
\end{aligned}
\qquad
\begin{aligned}
h(\mathbf{x}) &= (2\pi)^{-1/2} \\
g(\boldsymbol{\eta}) &= (-2\eta_2)^{1/2}\exp\left(\frac{\eta_1^2}{4\eta_2}\right)
\end{aligned}
$$