「パターン認識と機械学習」
2-2-1 ディリクレ分布

新納浩幸

多項分布

試行の結果が A1 ～AK のどれか、
Ak が起こる確率が μ_k

$$\mu_1 + \mu_2 + \cdots + \mu_K = 1$$

N 回の試行で Ak が m_k 回起こる確率

$$m_1 + m_2 + \cdots + m_K = N$$

$$Mult(m_1, m_2, \cdots, m_K) = \frac{N!}{m_1! m_2! \cdots m_K!} \prod_{k=1}^{K} \mu_k^{m_k}$$

多項分布の事前分布

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_K)$$

↓

各　μ_k　を確率変数と考える

K 次元の確率変数となる

事前分布

ディリクレ分布

多項分布の事前分布として標準的に利用される分布

$$Dir(\boldsymbol{\mu} \mid \boldsymbol{\alpha}) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\cdots\Gamma(\alpha_K)} \prod_{k=1}^{K} \mu_k^{\alpha_k - 1}$$

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_K)$$

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

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_K)$$

共役事前分布

なぜ多項分布の事前分布にディリクレ分布を使うか？

ディリクレ分布は多項分布の共役事前分布

事後分布が事前分布と同じ分布になるような
事前分布、、、計算が楽になる

事後分布

$$\frac{\Gamma(\alpha_0+N)}{\Gamma(\alpha_1+m_1)\Gamma(\alpha_2+m_2)\cdots\Gamma(\alpha_K+m_k)}\prod_{k=1}^{K}\mu_k^{\alpha_k+m_k-1}$$