「パターン認識と機械学習」
2-4-1 最尤推定と十分推定量

新納浩幸

ηの最尤推定

x 上の指数分布族:

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

x で積分して正規化する

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

積分の外に出ることに注意

η について両辺を微分

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

$$+\,g(\boldsymbol{\eta})\int h(\mathbf{x})\exp\{\boldsymbol{\eta}^T u(\mathbf{x})\}u(\mathbf{x})d\mathbf{x}=0$$

$$-\frac{\nabla g(\boldsymbol{\eta})}{g(\boldsymbol{\eta})}=g(\boldsymbol{\eta})\int h(\mathbf{x})\exp\{\boldsymbol{\eta}^T u(\mathbf{x})\}u(\mathbf{x})d\mathbf{x}$$

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

注意

$E(u(\mathbf{x}))$ は $g(\boldsymbol{\eta})$ のみに依存

$$-\nabla\ln g(\boldsymbol{\eta})=E(u(\mathbf{x}))$$

$\mathbf{X}=\{\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_N\}$ に対する尤度関数

$$p(\mathbf{X}|\boldsymbol{\eta})=\left(\prod_{n=1}^{N}h(\mathbf{x}_n)\right)g(\boldsymbol{\eta})^N\exp\left\{\boldsymbol{\eta}^T\sum_{n=1}^{N}u(\mathbf{x}_n)\right\}$$

両辺の対数を取る

$$\ln p(\mathbf{X}|\boldsymbol{\eta})=\sum_{n=1}^{N}h(\mathbf{x}_n)+N\ln g(\boldsymbol{\eta})+\boldsymbol{\eta}^T\sum_{n=1}^{N}u(\mathbf{x}_n)$$

微分して 0 の極値条件を使って

$$-\nabla\ln g(\boldsymbol{\eta}_{ML})=\frac{1}{N}\sum_{n=1}^{N}u(\mathbf{x}_n)$$

十分統計量

$$-\nabla \ln g(\boldsymbol{\eta}_{ML}) = \frac{1}{N}\sum_{n=1}^{N} u(\mathbf{x}_n)$$

$\boldsymbol{\eta}_{ML}$ は $\sum_{n=1}^{N} u(\mathbf{x}_n)$ のみに依存

十分統計量

十分統計量の例

ベルヌーイ分布(01分布)

$$u(x) = x$$ なので $$\sum_n x_n$$ で十分

ガウス分布(正規分布)

$$u(x) = (x, x^2)^T$$ なので

$$\sum_n x_n$$ と $$\sum_n x_n^2$$ で十分