パターン認識と機械学習

２．３．４　ガウス分布の最尤推定

松本良太

ガウス分布の最尤推定

観測値 $\{x_n\}$ が、独立に得られたと仮定したデータ集合

$$\mathbf{X}=(x_1,\ldots x_N)$$

その分布のパラメータは最尤推定法で推定できる。

対数尤度関数は、以下のようになる。

$$\ln p(\mathbf{X}\mid\boldsymbol{\mu},\boldsymbol{\Sigma})=-\frac{ND}{2}\ln(2\pi)-\frac{N}{2}\ln|\Sigma|$$

$$-\frac{1}{2}\sum_{n=1}^{N}(x_n-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(x_n-\boldsymbol{\mu})$$

ガウス分布の最尤推定

これを整理すると、尤度関数は次の2つの量によってのみデータ集合に依存していることがわかる。

$$\sum_{n=1}^{N} x_n, \sum_{n=1}^{N} x_n x_n^T$$

これらをガウス分布の***十分統計量***という。

ガウス分布の最尤推定

対数尤度のμについての導関数は

$$\frac{\partial}{\partial \boldsymbol{\mu}} \ln p(\mathbf{X} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^{N} \boldsymbol{\Sigma}^{-1}(x_n - \boldsymbol{\mu})$$

で与えられる。この導関数を0とおくと、最尤推定による平均についての解が得られる。

$$\boldsymbol{\mu}_{ML} = \frac{1}{N}\sum_{n=1}^{N} x_n$$

これは、観測されたデータ集合の平均である。

ガウス分布の最尤推定

・Σについての最大化

$$\boldsymbol{\Sigma}_{ML} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \boldsymbol{\mu}_{ML})(x_n - \boldsymbol{\mu}_{ML})^T$$

μ MLが含まれているのは、この結果がμとΣについて同時に最大化したものであるため。
この***μ ML***は***Σ ML***に依存しないので、まず***μ ML***を求めてから、***Σ ML***を求めることができる。

ガウス分布の最尤推定

真の分布の元での最尤推定解の期待値を評価すると、次の結果を得る。

$$\mathbf{E}[\boldsymbol{\mu}_{ML}] = \boldsymbol{\mu}$$

$$\mathbf{E}[\boldsymbol{\Sigma}_{ML}] = \frac{N-1}{N}\boldsymbol{\Sigma}$$

平均についての最尤推定量の期待値は真の平均に等しい。

ガウス分布の最尤推定

共分散の最尤推定量の期待値は真の値より小さく、
不偏推定にならない。
↓
次式のような別の推定量を定義すると補正できる。

$$\tilde{\boldsymbol{\Sigma}} = \frac{1}{N-1}\sum_{n=1}^{N}(x_n - \boldsymbol{\mu}_{ML})(x_n - \boldsymbol{\mu}_{ML})^T$$

この期待値はΣに等しくなる。