パターン認識と機械学習（上）
２．３．１条件付きガウス分布

２００８年6月13日

田中洸一

内容

多変量ガウス分布について、2つの変数集合の同時分布がガウス分布に従う

↓

一方の変数集合が与えられたときの、もう一方の集合の条件付きガウス分布もガウス分布

条件付きガウス分布の表現とその平均と共分散について考える

定義(1)

- ガウス分布N(x | μ , Σ)に従うD次元ベクトルをxとし、以下のように定義する。

$$x = \begin{pmatrix} x_a \\ x_b \end{pmatrix} \quad (2.65) \qquad \text{平均} \quad \mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} \quad (2.66)$$

共分散 $$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \quad (2.67)$$

xを互いに素な部分集合x_aとx_bに分割

定義（2）

- 共分散の逆行列Λについても定義する。

$$\Sigma^{-1} \equiv \Lambda = \begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix} \quad (2.68)$$

- 精度行列といい、後に表す式が簡素になるため用いる。

条件付き分布の表現方法（1）

x_bを固定し、得られた式をx_a上の確率となるように正規化する

↓

同時分布p(x)=p（x_a,x_b）で条件付き分布を求めることができる

↓

だが、

条件付き分布の表現方法（2）

その正規化の代わりに、ガウス分布の指数部分の二次形式（※）について考え、正規化係数を求めたほうが効率がいい

$$※\quad \Delta^2 = (x-\mu)^{\mathrm{T}} \Sigma^{-1} (x-\mu) \tag{2.44}$$

先ほどの二次形式に、定義したx・平均・精度を対応させる

$$-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)=$$

$$-\frac{1}{2}(x_a-\mu_a)^{\mathrm{T}}\Lambda_{aa}(x_a-\mu_a)-\frac{1}{2}(x_a-\mu_a)^{\mathrm{T}}\Lambda_{ab}(x_b-\mu_b)$$

$$-\frac{1}{2}(x_b-\mu_b)^{\mathrm{T}}\Lambda_{ba}(x_a-\mu_a)-\frac{1}{2}(x_b-\mu_b)^{\mathrm{T}}\Lambda_{bb}(x_b-\mu_b) \tag{2.70}$$

x_aの関数と見ると二次形式となっているので、条件付き分布もガウス分布となる

平均$\mu_{a|b}$と分散$\Sigma_{a|b}$を求める
ガウス分布を次のように表す

$$-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu) = -\frac{1}{2}x^{\mathrm{T}}\Sigma^{-1}x + x^{\mathrm{T}}\Sigma^{-1}\mu + const$$

(2.71)　※constはxとは独立な項

つまり、上の式の右辺の形式に表現すれば、

- xの二次の項の係数行列は<u>逆共分散Σ^{-1}</u>と等しい
- xの線形項の係数と<u>$\Sigma^{-1}\mu$</u>が等しい

p(x_a|x_b)の分散

(2.70)のx_aの２次の項を取り出すと

$$-\frac{1}{2}x_a^{\mathrm{T}}\Lambda_{aa}x_a$$

xの二次の項の係数行列は逆共分散Σ^{-1}と等しいので

$$\Sigma_{a|b} = \Lambda_{aa}^{-1} \qquad (2.73)$$

が得られる

$p(x_a|x_b)$の平均

(2.70)のx_aの線形の項を取り出すと

$$x_a^{\mathrm{T}}\left\{\Lambda_{aa}\mu_a-\Lambda_{ab}\left(x_b-\mu_b\right)\right\}$$

xの線形項の係数と$\underline{\Sigma^{-1}\mu}$が等しいので

$$\begin{aligned}\mu_{a|b}&=\Sigma_{a|b}\left\{\Lambda_{aa}\mu_a-\Lambda_{ab}\left(x_b-\mu_b\right)\right\}\\&=\mu_a-\Lambda_{aa}^{-1}\Lambda_{ab}\left(x_b-\mu_b\right)\quad(\because\ \Sigma_{a|b}=\Lambda_{aa}^{-1})\end{aligned}$$

(2.75)

が得られる

共分散行列による平均・分散（1）

$p(x_a|x_b)$の平均・分散を分割された共分散行列
で表すこともできる

分割された精度行列を分割された共分散行列で表す

$$\Lambda_{aa} = \left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1} \qquad (2.79)$$

$$\Lambda_{ab} = -\left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1}\Sigma_{ab}\Sigma_{bb}^{-1} \qquad (2.80)$$

共分散行列による平均・分散（2）

(2.79)と(2.80)を分割された精度行列で表した

平均・分散に代入{(2.73)と(2.75)に代入}

平均 $$\mu_{a|b} = \mu_b + \Sigma_{ab}\Sigma_{bb}^{-1}\left(x_b - \mu_b\right) \qquad (2.81)$$

分散 $$\Sigma_{a|b} = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba} \qquad (2.82)$$

※ $\Sigma_{a|b} = \Lambda_{aa}^{-1}$ (2.73)

<u>分散については、共分散行列で表すよりも</u>

<u>精度行列で表したほうが簡素に表せる</u>