パターン認識と機械学習

2.3.1 条件付きガウス分布

吉田拓夢

多変量ガウス分布の性質

2つの変数集合の同時分布がガウス分布に従う

変数集合の一方が与えられた時のもう一方の
条件付き分布もガウス分布になる

(周辺分布についても同様)

条件付きガウス分布

- $\mathbf{x}$はガウス分布$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma})$に従うD次元ベクトル

- xを2つの互いに素な部分集合に分割

→対応する平均ベクトルμ、共分散行列Σも同様

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \end{pmatrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{pmatrix} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{pmatrix}$$

共分散行列の対称性 $\boldsymbol{\Sigma}^{\mathrm{T}} = \boldsymbol{\Sigma}$ より$\boldsymbol{\Sigma}_{aa}$と $\boldsymbol{\Sigma}_{bb}$
も対称.また$\boldsymbol{\Sigma}_{ba} = \boldsymbol{\Sigma}_{ab}^{\mathrm{T}}$

条件付きガウス分布

- 精度行列(precision matrix)

→共分散の逆行列

$$\mathbf{\Lambda} \equiv \mathbf{\Sigma}^{-1}$$

- 精度行列についても分割

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{\Lambda}_{aa} & \mathbf{\Lambda}_{ab} \\ \mathbf{\Lambda}_{ba} & \mathbf{\Lambda}_{bb} \end{pmatrix}$$

条件付きガウス分布

- 条件付き分布$p(x_a|x_b)$

  x_bを観測された値で固定.
  得られた式をx_a上の正当な確率になるよう**正規化**.

ガウス分布の指数部分の二次形式を考え、
計算の最後に正規化係数を求める.

条件付きガウス分布

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

$$-\frac{1}{2}(\mathbf{x}_a-\boldsymbol{\mu}_a)^{\mathrm{T}}\Lambda_{aa}(\mathbf{x}_a-\boldsymbol{\mu}_a)-\frac{1}{2}(\mathbf{x}_a-\boldsymbol{\mu}_a)^{\mathrm{T}}\Lambda_{ab}(\mathbf{x}_b-\boldsymbol{\mu}_b)$$

$$-\frac{1}{2}(\mathbf{x}_b-\boldsymbol{\mu}_b)^{\mathrm{T}}\Lambda_{ba}(\mathbf{x}_a-\boldsymbol{\mu}_a)-\frac{1}{2}(\mathbf{x}_b-\boldsymbol{\mu}_b)^{\mathrm{T}}\Lambda_{bb}(\mathbf{x}_b-\boldsymbol{\mu}_b)$$

↑xaの関数と見ると二次形式なので、対応する条件付き分布p(xa|xb)もガウス分布と分かる

平均と共分散

↓一般のガウス分布 $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma})$ の指数部分

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

- xの二次の項の係数行列→逆共分散行列 $\boldsymbol{\Sigma}^{-1}$
- xの線形項の係数→ $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$

平均と共分散

- 条件付きガウス分布$p(x_a|x_b)$→平均$\mu_{a|b}$,共分散$\Sigma_{a|b}$

xbを定数としてxaの関数的依存性を考える.
xaについての二次の項をすべて取り出すと

$$-\frac{1}{2}\mathbf{x}_a^{\mathrm{T}}\mathbf{\Lambda}_{aa}\mathbf{x}_a$$

上記より

$$\mathbf{\Sigma}_{a|b} = \mathbf{\Lambda}_{aa}^{-1}$$

平均と共分散

xaについて線形の項をすべて考えると

$$\mathbf{x}_a^{\mathrm{T}}\{\boldsymbol{\Lambda}_{aa}\boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{ab}(\mathbf{x}_b - \boldsymbol{\mu}_b)\}$$

↑xaの係数は $\boldsymbol{\Sigma}_{a|b}^{-1}\boldsymbol{\mu}_{a|b}$ と等しくなるので

$$\begin{aligned}\boldsymbol{\mu}_{a|b} &= \boldsymbol{\Sigma}_{a|b}\{\boldsymbol{\Lambda}_{aa}\boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{ab}(\mathbf{x}_b - \boldsymbol{\mu}_b)\} \\ &= \boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{aa}^{-1}\boldsymbol{\Lambda}_{ab}(\mathbf{x}_b - \boldsymbol{\mu}_b)\end{aligned}$$

- 平均$\mu_{a|b}$,共分散$\Sigma_{a|b}$は分割された精度行列で表現

平均と共分散

- 分割された共分散行列で表現する
  (以下の式を利用)

$$\begin{pmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{\Lambda}_{aa} & \boldsymbol{\Lambda}_{ab} \\ \boldsymbol{\Lambda}_{ba} & \boldsymbol{\Lambda}_{bb} \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{M} & -\mathbf{M}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{M} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{M}\mathbf{B}\mathbf{D}^{-1} \end{pmatrix}$$

$$\Rightarrow \quad \mathbf{M} = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$$

$\mathbf{M}^{-1}$ は部分行列Dに対するシューア補行列

平均と共分散

$$\mathbf{\Lambda}_{aa} = (\mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba})^{-1}$$

$$\mathbf{\Lambda}_{ab} = -(\mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba})^{-1}\mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}$$

上記の式が得られ、以下が求まる

$$\boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a + \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}(\mathbf{x}_b - \boldsymbol{\mu}_b)$$ ←xbの線形関数

$$\mathbf{\Sigma}_{a|b} = \mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba}$$ ←xbとは独立

⇒ 線形ガウスモデル