06t4027h 斎藤高行

9章　ベイズ推定法

9.1 ベイズ推定法の定義

9.2 ベイズ推定法と最尤推定の違い

確率の定義

訓練標本を $\chi := \{x_i\}_{i=1}^{n}$ とする。
パラメータ θ を確率的な変数とみなすとき、
次のような確率がそれぞれ定義できる

$p(\theta \mid \chi)$: θ の事後確率
$p(\theta)$: θ の事前確率
$p(\chi \mid \theta)$: θ の尤度
$p(\chi, \theta)$: χ と θ の同時確率

θ の尤度

- θ の関数とみなし、 $p(\chi \mid \theta)$ で定義する
- $p(\chi \mid \theta) = \prod_{i=1}^{n} q(x_i \mid \theta)$ で与えられる

※ θ は確率変数であり、モデルも $q(x_i \mid \theta)$ で与えられるとする

χとθの同時確率

- $p(\chi,\theta)=p(\chi\mid\theta)p(\theta)$ と表わすことができる

- θ に関して周辺化すると、周辺確率 $p(\chi)$ が得られる

$$p(\chi)=\int p(\chi,\theta)d\theta=\int p(\chi\mid\theta)p(\theta)d\theta$$

$$=\int\prod_{i=1}^{n}q(x_i\mid\theta)p(\theta)d\theta$$

ベイズ推定法の定義

- モデル $q(x_i \mid \theta)$ を事後確率 $p(\theta \mid \chi)$ に関して平均にすることで、$p(x)$ を推定する方法

- ベイズ推定法による確率密度関数の推定量 $\hat{p}Bayes(x)$ は次のようである

$$\hat{p}Bayes(x) := \int q(x \mid \theta) p(\theta \mid \chi) d\theta$$

ベイズ推定法の定義

- θ の事後確率 $p(\theta \mid \chi)$ は、次のように表せる

$$p(\theta \mid \chi) = \frac{p(\chi \mid \theta)p(\theta)}{p(\chi)} = \frac{\prod_{i=1}^{n} q(x_i \mid \theta)p(\theta)}{\int \prod_{i=1}^{n} q(x_i \mid \theta')p(\theta')d\theta'}$$

ベイズ推定法の定義

- したがって、推定量 $\hat{p}Bayes(x)$ は

$$\hat{p}Bayes(x) = \frac{\int q(x \mid \theta) \prod_{i=1}^{n} q(x_i \mid \theta) p(\theta) d\theta}{\int \prod_{i=1}^{n} q(x_i \mid \theta') p(\theta') d\theta'}$$

最尤推定法との違い

確率密度関数の近似の方法に差異がある

- 最尤推定法:代表パラメータ $\hat{\theta}_{ML}$ により近似
- ベイズ推定法:無数のパラメータを事後確率 $p(\theta \mid \chi)$ に従って平均することにより近似

モデル誤りとベイズ推定法

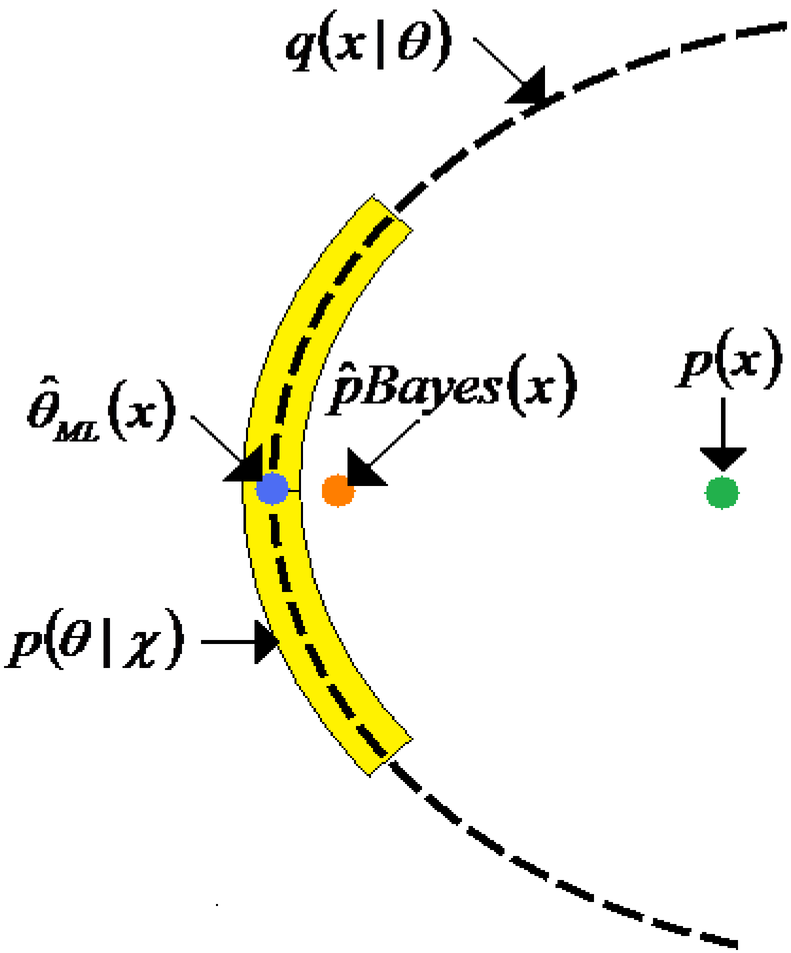
<u>モデル誤り</u>

推定したい真の確率密度関数 $p(x)$ がパラメトリックモデルに含まれていない状況のこと

モデル誤りがおきているときに、ベイズ推定法で近似することにより、真の確率密度関数に近づくことがある

(次のスライドに図を載せる)

最尤推定法との違い



・ベイズ推定法での推定量 $\hat{p}Bayes(x)$ は、一般にパラメトリック モデルの外にはみ出す。

・図の $p(x)$ は、モデル誤りの例を示したものである。