Laplace Approximation

In the context of Bayes factors as an approach to Bayesian model comparison

STAT 206 5 Mar 21

DD AM discussion section

Remember Bayes factors from week 1?

Suppose that we have uncovered (1) an ensemble $\mathcal{M} = \{ M_1, M_2, \dots, M_m \}$ of worthwhile-to-consider models (e.g., using the cheating approach)

$\mathbb{P} = (\mathcal{Q}, \mathcal{C}) \rightarrow (\underset{\sim}{\theta}, \mathcal{D}, \mathcal{B})$ ← unknown quantities of principal interest

$\rightarrow M_j$: $(j = 1, \dots, m)$

$$\begin{cases} (\underset{\sim}{\theta}_j \mid [PM_j]\, \mathcal{B}) \sim p(\underset{\sim}{\theta}_j \mid [PM_j]\, \mathcal{B}) \\ (Y_i \mid \underset{\sim}{\theta}_j\, [SM_j]\, \mathcal{B}) \overset{IID}{\sim} p(y_i \mid \underset{\sim}{\theta}_j\, [SM_j]\, \mathcal{B}) \end{cases}$$

$(i = 1, \dots, n)$

$$\mathcal{D} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

ex. NB10 case study $\mathcal{M} = \{ \text{Gaussian}, t \}$ $(m = 2)$

$$M_1 : \begin{cases} \underset{\sim}{\theta}_N = (\mu, \sigma_N) \quad (\underset{\sim}{\theta}_1 \mid [PM_1]\, \mathcal{B}) \sim p(\underset{\sim}{\theta}_1 \mid [PM_1]\, \mathcal{B}) \\ (Y_i \mid \underset{\sim}{\theta}_1\, [SM_1 : N]\, \mathcal{B}) \overset{IID}{\sim} p(y_i \mid \underset{\sim}{\theta}_1\, [SM_1 : N]\, \mathcal{B}) \end{cases}$$

(2)

$$M_2: \begin{cases} (\underset{\sim}{\theta}_2 \mid [PM_2] \, \mathcal{B}) \sim p(\underset{\sim}{\theta}_2 \mid [PM_2] \, \mathcal{B}) \\ (Y_i \mid \underset{\sim}{\theta}_2 \, [SM_2 : \mathbb{F}] \, \mathcal{B}) \sim p(y_i \mid \underset{\sim}{\theta}_2 \, [SM_2 : \mathbb{F}] \, \mathcal{B}) \end{cases}$$

$(i = 1, \dots, n)$ $\underset{\sim}{\theta}_2 = (\mu, \sigma_t, \nu)$ not the same as σ_N in M_1

~~here~~

$\theta = \mu$ has the same meaning in both models

Since $|\mathcal{M}| = m < \infty$ (i.e., the number of models in $\mathcal{M}$ is finite) it's sufficient to have a method that compares models 2 at a time: M_1 vs. M_2; better of (M_1, M_2) vs. M_3; best M so far vs. M_4; ...

The Bayes factor story makes the following provisional (& heroic) assumptions: (1) There is a "true" data-generating model M_{DG}; (2) $M_{DG} \in \mathcal{M}$.

(this is called the $\underline{\mathcal{M}\text{-closed}}$ viewpoint)

Then first pretend that $\mathcal{M} = \{M_1, M_2\}$; then

$$\left[\frac{p(M_2 \mid D\, B)}{p(M_1 \mid D\, B)}\right] = \left[\frac{p(M_2 \mid B)}{p(M_1 \mid B)}\right]\left[\frac{p(D \mid M_2\, B)}{p(D \mid M_1\, B)}\right] \quad (*)$$

(posterior odds in favor of M_2 over M_1 given D & B) = (prior odds in favor of M_2 over M_1 given B) · (Bayes factor in favor of M_2 over M_1 given D and B)

Q: How specify prior model probabilities $p(M_j \mid B)$?

A: "Please ask me an easier question!"

Seriously: many people try to "avoid" this issue by assigning a uniform prior across models:

$$p(M_j \mid B) = \begin{cases} \frac{1}{m} & \text{for } j = 1, \dots, m \text{ in } \mathcal{M} = \{M_1, \dots, M_m\} \\ 0 & \text{else} \end{cases}$$