Формулы

September 2016

Рассмотрим следующуюу вероятностную модель:

$$p(T, \theta|\alpha, X) = p(T|X, \theta)p(\theta|\alpha) \tag{1}$$

Дана обучающая выборка (T_{tr}, X_{tr}).

$$p(T_{tr}|X_{tr}, \alpha) = \int p(T_{tr}, \theta|X_{tr}, \alpha)d\theta = \{ \text{ используя определение модели (1)} \} =$$

$$= \int p(T_{tr}|X_{tr}, \theta)p(\theta|\alpha)d\theta \to \max_{\alpha}$$

Предположим, что максимум достигается на α_{max}.

Вычисление апостериорного распределения:

$$p(\theta|X_{tr}, T_{tr}, \alpha_{\max}) = \{ \text{ используя формулу Байеса } \} = \frac{p(\theta, T_{tr}|X_{tr}, \alpha_{\max})}{p(T_{tr}|X_{tr}, \alpha_{\max})} =$$

$$= \{ \text{ используя определение модели (1) } \} = \frac{p(T_{tr}|X_{tr}, \theta)p(\theta|\alpha_{\max})}{\int p(T_{tr}, \theta|X_{tr}, \alpha_{\max})d\theta} = \frac{p(T_{tr}|X_{tr}, \theta)p(\theta|\alpha_{\max})}{\int p(T_{tr}|X_{tr}, \theta)p(\theta|\alpha_{\max})d\theta}$$

Пусть дана тестовая выборка X_{ts} хотим вычислить распределение на T_{ts}. Предположение о том, что выборка (T_{ts}, X_{ts}) и выборка (T_{tr}, X_{tr}) независимы формализуется следующем образом:

$$p(T_{tr}, T_{ts}|X_{ts}, X_{tr}, \theta) = p(T_{ts}|X_{ts}, \theta)p(T_{tr}|X_{tr}, \theta) \tag{2}$$

По аналогии, как мы пишем, что если выборка $\{(x_1, y_1), ..., (x_l, y_l)\}$ – i.i.d., то $p(y_1, ..., y_l|x_1, ..., x_l, \theta) = \prod_{i=1}^{l} p(y_i|x_i, \theta)$.

Покажем, что

$$p(T_{tr}|X_{ts}, X_{tr}, \alpha_{\max}) = p(T_{tr}|X_{tr}, \alpha_{\max}) \tag{3}$$

$$p(T_{tr}|X_{tr}, X_{ts}, \alpha_{\max}) = \int p(T_{ts}, T_{tr}|X_{tr}, X_{ts}, \alpha_{\max})dT_{ts} = \int_{T_{ts}} \int_{\theta} p(\underbrace{T_{ts}, T_{tr}}_{T}, \theta|\underbrace{X_{ts}, X_{tr}}_{X}, \alpha_{\max})dT_{ts}d\theta =$$

$$= \{\text{в силу задания модели (1)}\} = \int_{T_{ts}} \int_\theta p(T_{ts}, T_{tr}|X_{ts}, X_{tr}, \theta) p(\theta|\alpha_{\max}) dT_{ts} d\theta =$$

$$= \{\text{в силу (2)}\} = \int_{T_{ts}} \int_\theta p(T_{ts}|X_{ts}, \theta) p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha_{\max}) dT_{ts} d\theta =$$

$$= \int_\theta \underbrace{\left(\int_{T_{ts}} p(T_{ts}|X_{ts}, \theta) dT_{ts} \right)}_{=1} p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha_{\max}) d\theta = \int_\theta p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha_{\max}) d\theta =$$

$$= \int_\theta p(T_{tr}, \theta|X_{tr}, \alpha_{\max}) d\theta = p(T_{tr}|X_{tr}, \alpha_{\max})$$

Оценим распределение на T_{ts}

$$p(T_{ts}|X_{ts}, X_{tr}, T_{tr}, \alpha_{\max}) = \frac{p(T_{ts}, T_{tr}|X_{ts}, X_{tr}, \alpha_{\max})}{p(T_{tr}|X_{tr}, X_{ts}, \alpha_{\max})} = \{\text{в силу (3)}\} = \frac{p(T_{ts}, T_{tr}|X_{ts}, X_{tr}, \alpha_{\max})}{p(T_{tr}|X_{tr}, \alpha_{\max})} =$$

$$= \frac{1}{p(T_{tr}|X_{tr}, \alpha_{\max})} \int_\theta p(\underbrace{T_{ts}, T_{tr}}_{T}, \theta| \underbrace{X_{ts}, X_{tr}}_{X}, \alpha_{\max}) d\theta =$$

$$= \{\text{в силу задания модели (1)}\} = \frac{1}{p(T_{tr}|X_{tr}, \alpha_{\max})} \int_\theta p(T_{ts}, T_{tr}|X_{ts}, X_{tr}, \theta) p(\theta|\alpha_{max}) d\theta =$$

$$= \int p(T_{ts}|X_{ts}, \theta) \underbrace{\frac{p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha_{\max})}{p(T_{tr}|X_{tr}, \alpha_{\max})}}_{p(\theta|T_{tr}, X_{tr}, \alpha_{\max})} d\theta = \int_\theta p(T_{ts}|X_{ts}, \theta) p(\theta|T_{tr}, X_{tr}, \alpha_{\max}) d\theta$$

Если рассматривается модель

$$p(T, \theta, \alpha|X) = p(T|X, \theta) p(\theta|\alpha) p(\alpha) \qquad (4)$$

Тогда необходимо вычислить

$$p(\alpha|X_{tr}, T_{tr}) = \frac{p(T_{tr}, \alpha|X_{tr})}{p(T_{tr}|X_{tr})}$$

Вычислим числитель и знаменатель

$$p(T_{tr}, \alpha|X_{tr}) = \int_\theta p(T_{tr}, \theta, \alpha|X_{tr}) d\theta = \int_\theta p(T_{tr}, \theta, \alpha|X_{tr}) d\theta = \int_\theta p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha) p(\alpha) d\theta$$

$$p(T_{tr}|X_{tr}) = \int_\alpha \int_\theta p(T_{tr}, \theta, \alpha|X_{tr}) d\theta d\alpha = \int_\alpha \int_\theta p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha) p(\alpha) d\theta d\alpha$$

Покажем, что

$$p(\alpha|X_{ts}, X_{tr}, T_{tr}) = p(\alpha|X_{tr}, T_{tr}) \tag{5}$$

$$p(\alpha|X_{tr}, X_{ts}, T_{tr}) = \frac{p(\alpha, T_{tr}|X_{tr}, X_{ts})}{p(T_{tr}|X_{tr}, X_{ts})} = \frac{p(\alpha, T_{tr}|X_{tr}, X_{ts})}{\int_\alpha p(T_{tr}, \alpha|X_{tr}, X_{ts})d\alpha} =$$

$$= \{\text{ниже показано, что } p(\alpha, T_{tr}|X_{tr}, X_{ts}) = p(\alpha, T_{tr}|X_{tr})\} = \frac{p(\alpha, T_{tr}|X_{tr})}{\int_\alpha p(\alpha, T_{tr}|X_{tr})d\alpha} =$$

$$= \frac{p(\alpha, T_{tr}|X_{tr})}{p(T_{tr}|X_{tr})} = p(\alpha|X_{tr}, T_{tr})$$

$$p(T_{tr}, \alpha|X_{ts}, X_{tr}) = \int_{T_{ts}} \int_\theta p(\alpha, \theta, \underbrace{T_{tr}, T_{ts}}_{T} | \underbrace{X_{ts}, X_{tr}}_{X}) d\theta dT_{ts} =$$

$$= \{\text{в силу определения модели (4)}\} = \int_{T_{ts}} \int_\theta p(T_{tr}, T_{ts}|X_{ts}, X_{tr}, \theta) p(\theta|\alpha) p(\alpha) d\theta dT_{ts} =$$

$$= \{\text{в силу (2)}\} = \int_{T_{ts}} \int_\theta p(T_{ts}|X_{ts}, \theta) p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha) p(\alpha) d\theta dT_{ts} =$$

$$= \int_\theta \underbrace{\left(\int_{T_{ts}} p(T_{ts}|X_{ts}, \theta) dT_{ts}\right)}_{=1} p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha) p(\alpha) d\theta = \int_\theta p(T_{tr}|X_{tr}, \theta) p(\theta|\alpha) p(\alpha) d\theta =$$

$$= \int_\theta p(T_{tr}, \theta, \alpha|X_{tr}) d\theta = p(T_{tr}, \alpha|X_{tr})$$

Теперь мы можем оценить распределение T_{ts} для модели 4.

$$p(T_{ts}|X_{ts}, X_{tr}, T_{tr}) = \int_\alpha p(T_{ts}, \alpha|X_{ts}, T_{tr}, X_{tr}) d\alpha = \int_\alpha p(T_{ts}|X_{ts}, X_{tr}, T_{tr}, \alpha) p(\alpha|X_{ts}, X_{tr}, T_{tr}) d\alpha =$$

$$= \{\text{в силу (5)}\} = \int_\alpha p(T_{ts}|X_{ts}, X_{tr}, T_{tr}, \alpha) p(\alpha|X_{tr}, T_{tr}) d\alpha$$