ВМК МГУ

Семинар 5. Метод релевантных векторов

Курс: Байесовские методы в машинном обучении, 2016

1. Рассматривается модель регрессии релевантных векторов. Можно показать, что для логарифма обоснованности справедлива следующая нижняя оценка:

$$\log p(\boldsymbol{t}|X, A, \beta) = \log \int \mathcal{N}(\boldsymbol{t}|X\boldsymbol{w}, \beta^{-1}I)\mathcal{N}(\boldsymbol{w}|\boldsymbol{0}, A)d\boldsymbol{w} \geq$$
$$\geq \frac{N}{2}\log\beta - \frac{N}{2}\log 2\pi + \frac{1}{2}\log\det A - \frac{\beta}{2}\|\boldsymbol{t} - X\boldsymbol{\mu}\|^2 - \frac{1}{2}\boldsymbol{\mu}^T A\boldsymbol{\mu} - \frac{1}{2}\log\det(\beta X^T X + A).$$

   Найти формулы для $\boldsymbol{\mu}, A = \mathrm{diag}(\alpha_1, \ldots, \alpha_d), \beta$, соответствующие покомпонентной максимизации данной нижней оценки.

2. Доказать, что прогнозное распределение в регрессии релевантных векторов имеет вид

$$p(t_{new}|\boldsymbol{x}_{new}, \boldsymbol{t}, X, A_{ME}, \beta_{ME}) = \int \mathcal{N}(t_{new}|\boldsymbol{w}^T\boldsymbol{x}_{new}, \beta_{ME}^{-1})p(\boldsymbol{w}|\boldsymbol{t}, X, A_{ME}, \beta_{ME})d\boldsymbol{w} =$$
$$= \mathcal{N}(t_{new}|\boldsymbol{x}_{new}^T\boldsymbol{\mu}, \beta_{ME}^{-1} + \boldsymbol{x}_{new}^T\Sigma\boldsymbol{x}_{new}),$$
$$\Sigma = (\beta_{ME}X^TX + A_{ME})^{-1},$$
$$\boldsymbol{\mu} = \Sigma\beta_{ME}X^T\boldsymbol{t}.$$

3. Найти формулы для обучения регрессии релевантных векторов в случае $A = \alpha I$.

4. Доказать оценку Бушара (Bouchard) для функции log-sum-exp:

$$\log(\sum_i \exp(y_i)) \leq \alpha + \sum_i \log(1 + \exp(y_i - \alpha)) \leq \alpha + \sum_i \left[\log(1 + \exp(\xi_i)) + \frac{y_i - \alpha - \xi_i}{2} + \frac{\tanh(\xi_i/2)}{4\xi_i}((y_i - \alpha)^2 - \xi_i^2)\right].$$