Часть 9: Нормализация

•••

Романов Михаил, Игорь Слинько

Метод максимального правдоподобия

$\{\vec{x}_i\}_{i=1}^{N}$

$P(\{x_i\}_{i=1}^{N}) = p(x_0) \cdot p(x_1) \cdot \ldots \cdot p(x_2) = \prod_{i=1}^{N} p(x_i)$

Метод максимального правдоподобия

$\{\vec{x}_i\}_{i=1}^N$

$P_w(\{x_i\}_{i=1}^N) = p_w(x_0) \cdot p_w(x_1) \cdot \ldots \cdot p_w(x_2) = \prod_{i=1}^N p_w(x_i) \longrightarrow \max_w$

$w^* = \arg\max_w P_w(\{x_i\}_{i=1}^N)$

Метод максимального правдоподобия

$\{\vec{x}_i\}_{i=1}^{N}$

$P_w(\{x_i\}_{i=1}^{N}) = p_w(x_0) \cdot p_w(x_1) \cdot \ldots \cdot p_w(x_2) = \prod_{i=1}^{N} p_w(x_i) \longrightarrow \max_w$

$w^* = \arg\max_w P_w(\{x_i\}_{i=1}^{N})$

$\arg\max_w P_w(\{x_i\}_{i=1}^{N}) = \arg\max_w \log P_w(\{x_i\}_{i=1}^{N})$

Метод максимального правдоподобия

$$\{\vec{x}_i\}_{i=1}^{N}$$

$$P_w(\{x_i\}_{i=1}^{N}) = p_w(x_0)\cdot p_w(x_1)\cdot\ldots\cdot p_w(x_2) = \prod_{i=1}^{N} p_w(x_i) \longrightarrow \max_w$$

$$w^* = \arg\max_w P_w(\{x_i\}_{i=1}^{N})$$

$$\arg\max_w P_w(\{x_i\}_{i=1}^{N}) = \arg\max_w \log P_w(\{x_i\}_{i=1}^{N})$$

$$\log P_w(\{x_i\}_{i=1}^{N}) = \log \prod_{i=1}^{N} p_w(x_i) = \sum_{i=1}^{N} \log p_w(x_i) \longrightarrow \max_w$$

Метод максимального правдоподобия

$$\log P_w(\{x_i\}_{i=1}^N) = \log \prod_{i=1}^N p_w(x_i) = \sum_{i=1}^N \log p_w(x_i) \longrightarrow \max_w$$

$$p_w(x_i) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad w = \{\mu\}$$

$$\log p_w(x_i) = -\log(\sqrt{2\pi}) - \log(\sigma) - \frac{(x-\mu)^2}{2\sigma^2}$$

$$\sum_{i=1}^N \log p_w(x_i) = -N\log(\sqrt{2\pi}) - N\log(\sigma) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^N (x_i-\mu)^2 \longrightarrow \max_{\{\mu\}}$$

$$\arg\max P_w = \arg\max \log P_w = \arg\max_{\{\mu\}} -\sum_{i=1}^N (x_i-\mu)^2 =$$

$$\arg\min_{\{\mu\}} \sum_{i=1}^N (x_i-\mu)^2$$

Метод максимального правдоподобия

$$\log P_w(\{x_i\}_{i=1}^N) = \log \prod_{i=1}^N p_w(x_i) = \sum_{i=1}^N \log p_w(x_i) \to \max_w$$

$$p_w(x_i) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad w = \{\mu\}$$

$$\log p_w(x_i) = -\log(\sqrt{2\pi}) - \log(\sigma) - \frac{(x-\mu)^2}{2\sigma^2}$$

$$\sum_{i=1}^N \log p_w(x_i) = -N\log(\sqrt{2\pi}) - N\log(\sigma) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^N (x_i-\mu)^2 \to \max_{\{\mu\}}$$

$$\operatorname{argmax} P_w = \operatorname{argmax} \log P_w = \operatorname*{argmax}_{\{\mu\}} -\sum_{i=1}^N (x_i-\mu)^2 = \operatorname*{argmin}_{\{\mu\}} \sum_{i=1}^N (x_i-\mu)^2$$

Выводим ВСЕ

1: heads

0: tails

N experiments

n heads

$P(\text{heads})$ – ?

heads: p tails: $(1-p)$ Sequence: 1011010110 ... 10

$$P(\text{Sequence}) = p^n(1-p)^{N-n}$$

$$P(n) = C_N^n \cdot p^n(1-p)^{N-n}; \quad C_N^n = \frac{N!}{n!(N-n)!}$$

$$\underset{p}{\operatorname{argmax}}\, P(n) = \underset{p}{\operatorname{argmax}}\, \log P(n)$$

Выводим ВСЕ

1: heads

0: tails

N experiments

n heads

p(heads) - ?

$$\arg\max_p P(n) = \arg\max_p \log P(n) =$$

$$\arg\max_p \log C_N^n + n \log p + (N-n)\log(1-p) = \arg\max_p \frac{n}{N}\log p + \left(1-\frac{n}{N}\right)\log(1-p)$$