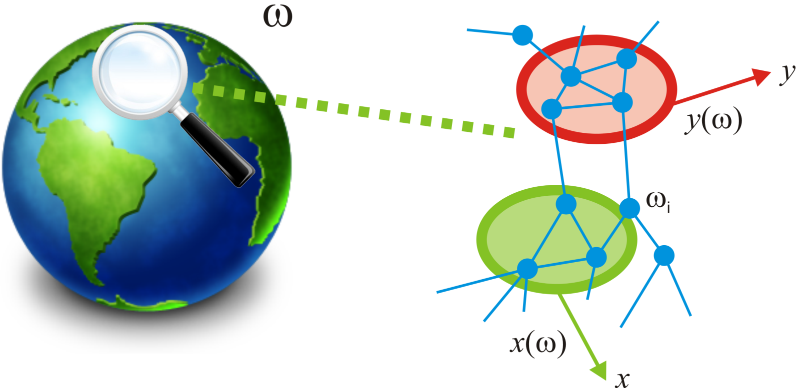
Wrocław University
of Science and Technology

Zadanie estymacji

Szymon Zaręba

Wartości opisujące rzeczywistość



Estymacja największej wiarygodności (1)

Danych jest N niezależnych realizacji $\mathcal{D} = \{\mathbf{x}_1 \ldots \mathbf{x}_N\}$ wektora losowego $\mathbf{x}$ o rozkładzie $p(\mathbf{x}|\theta)$.
Funkcja wiarygodności:

$$p(\mathcal{D}|\theta) = p(\mathbf{x}_1|\theta)p(\mathbf{x}_2|\theta)\cdots = \prod_{n=1}^{N} p(\mathbf{x}_n|\theta)$$

Zlogarytmowana funkcja wiarygodnosci:

$$\ln p(\mathcal{D}|\theta) = \sum_{n=1}^{N} \ln p(\mathbf{x}_n|\theta)$$

Estymatorem największej wiarygodności nazywamy θ_{ML}:

$$\theta_{ML} = \operatorname*{argmax}_{\theta} p(\mathcal{D}|\theta) = \operatorname*{argmax}_{\theta} \ln p(\mathcal{D}|\theta)$$

Estymacja największej wiarygodności (2)

1. Wyznaczyć funkcję wiarygodności
2. Wyznaczyć zlogarytmowaną funkcję wiarygodności
3. Wyznaczyć gradient zlogarytmowanej funkcji wiarygodności
4. Wyznaczyć ekstremum

Estymacja największej wiarygodności (3)

1. Wyznaczyć funkcję wiarygodności

Estymacja największej wiarygodności (3)

1. Wyznaczyć funkcję wiarygodności

$$p(\mathcal{D}|\theta) = \prod_{n=1}^{N} p(\mathbf{x}_n|\theta) = \prod_{n=1}^{N} \theta^{x_n}(1-\theta)^{1-x_n}$$

Estymacja największej wiarygodności (3)

1. Wyznaczyć funkcję wiarygodności

$$p(\mathcal{D}|\theta) = \prod_{n=1}^{N} p(\mathbf{x}_n|\theta) = \prod_{n=1}^{N} \theta^{x_n}(1-\theta)^{1-x_n}$$

2. Wyznaczyć zlogarytmowaną funkcję wiarygodności

Estymacja największej wiarygodności (3)

1. Wyznaczyć funkcję wiarygodności

$$p(\mathcal{D}|\theta) = \prod_{n=1}^{N} p(\mathbf{x}_n|\theta) = \prod_{n=1}^{N} \theta^{x_n}(1-\theta)^{1-x_n}$$

2. Wyznaczyć zlogarytmowaną funkcję wiarygodności

$$\ln p(\mathcal{D}|\theta) = \sum_{n=1}^{N} x_n \ln\theta + (1-x_n)\ln(1-\theta)$$

Estymacja największej wiarygodności (3)

1. Wyznaczyć funkcję wiarygodności

$$p(\mathcal{D}|\theta) = \prod_{n=1}^{N} p(\mathbf{x}_n|\theta) = \prod_{n=1}^{N} \theta^{x_n}(1-\theta)^{1-x_n}$$

2. Wyznaczyć zlogarytmowaną funkcję wiarygodności

$$\ln p(\mathcal{D}|\theta) = \sum_{n=1}^{N} x_n \ln\theta + (1-x_n)\ln(1-\theta)$$

3. Wyznaczyć gradient zlogarytmowanej funkcji wiarygodności

Estymacja największej wiarygodności (3)

1. Wyznaczyć funkcję wiarygodności

$$p(\mathcal{D}|\theta) = \prod_{n=1}^{N} p(\mathbf{x}_n|\theta) = \prod_{n=1}^{N} \theta^{x_n}(1-\theta)^{1-x_n}$$

2. Wyznaczyć zlogarytmowaną funkcję wiarygodności

$$\ln p(\mathcal{D}|\theta) = \sum_{n=1}^{N} x_n \ln\theta + (1-x_n)\ln(1-\theta)$$

3. Wyznaczyć gradient zlogarytmowanej funkcji wiarygodności

$$\frac{\partial \ln p(\mathcal{D}|\theta)}{\partial\theta} = \sum_{n=1}^{N} \frac{x_n}{\theta} - \frac{1-x_n}{1-\theta}$$

Estymacja największej wiarygodności (3)

1. Wyznaczyć funkcję wiarygodności

$$p(\mathcal{D}|\theta) = \prod_{n=1}^{N} p(\mathbf{x}_n|\theta) = \prod_{n=1}^{N} \theta^{x_n}(1-\theta)^{1-x_n}$$

2. Wyznaczyć zlogarytmowaną funkcję wiarygodności

$$\ln p(\mathcal{D}|\theta) = \sum_{n=1}^{N} x_n \ln\theta + (1-x_n)\ln(1-\theta)$$

3. Wyznaczyć gradient zlogarytmowanej funkcji wiarygodności

$$\frac{\partial \ln p(\mathcal{D}|\theta)}{\partial\theta} = \sum_{n=1}^{N} \frac{x_n}{\theta} - \frac{1-x_n}{1-\theta}$$

4. Wyznaczyć ekstremum

Estymacja maksymalnego a posteriori (1)

Danych jest N niezależnych realizacji $\mathcal{D} = \{\mathbf{x}_1 \ldots \mathbf{x}_N\}$ wektora losowego $\mathbf{x}$ o rozkładzie $p(\mathbf{x}|\theta)$ oraz rozkład *a priori* $p(\theta)$ parametru θ.

Estymatorem maksymalnego *a posteriori* (MAP) nazywamy θ_{MAP} maksymalizujący rozkład *a posteriori*:

$$\theta_{MAP} = \operatorname*{argmax}_{\theta} p(\theta|\mathcal{D})$$

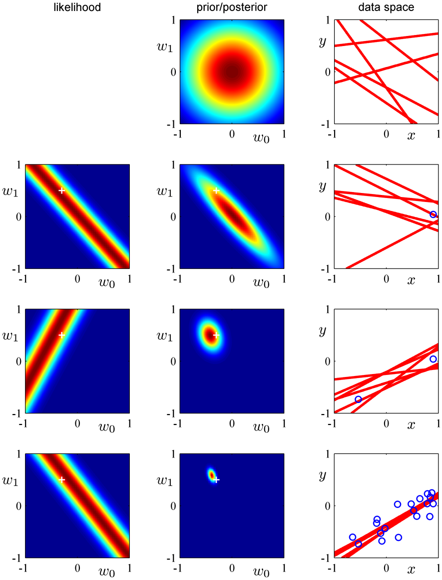
Jednocześnie korzystając ze wzoru Bayesa:

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})} = \frac{p(\theta) \prod_{n=1}^{N} p(x_n|\theta)}{p(\mathcal{D})}$$

Wtedy:

$$\theta_{MAP} = \operatorname*{argmax}_{\theta} \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})} = \operatorname*{argmax}_{\theta} p(\mathcal{D}|\theta)p(\theta)$$

Estymacja maksymalnego a posteriori (1)



Estymacja maksymalnego a posteriori (2)

1. Wyznaczyć rozkład a posteriori
2. Wyznaczyć zlogarytmowany rozkład a posteriori
3. Wyznaczyć gradient zlogarytmowanego rozkładu a posteriori
4. Wyznaczyć ekstremum

Estymacja maksymalnego a posteriori (3)

1. Wyznaczyć rozkład a posteriori:

Estymacja maksymalnego a posteriori (3)

1. Wyznaczyć rozkład a posteriori:

$$p(\theta|\mathcal{D}) = \frac{C\theta^{a-1}(1-\theta)^{b-1}\prod_{n=1}^{N}\theta^{x_n}(1-\theta)^{1-x_n}}{p(\mathcal{D})}$$

Estymacja maksymalnego a posteriori (3)

1. Wyznaczyć rozkład a posteriori:

$$p(\theta|\mathcal{D}) = \frac{C\theta^{a-1}(1-\theta)^{b-1}\prod_{n=1}^{N}\theta^{x_n}(1-\theta)^{1-x_n}}{p(\mathcal{D})}$$

2. Wyznaczyć zlogarytmowany rozkład a posteriori

Estymacja maksymalnego a posteriori (3)

1. Wyznaczyć rozkład a posteriori:

$$p(\theta|\mathcal{D}) = \frac{C\theta^{a-1}(1-\theta)^{b-1}\prod_{n=1}^{N}\theta^{x_n}(1-\theta)^{1-x_n}}{p(\mathcal{D})}$$

2. Wyznaczyć zlogarytmowany rozkład a posteriori

$$\begin{aligned}\ln p(\theta|\mathcal{D}) &= \ln C \\ &+ (a-1)\ln\theta + (b-1)\ln(1-\theta) \\ &+ \sum_{n=1}^{N} x_n \ln\theta + (1-x_n)\ln(1-\theta) \\ &- \ln p(\mathcal{D})\end{aligned}$$

Estymacja maksymalnego a posteriori (3)

3. Wyznaczyć gradient zlogarytmowanego rozkładu a posteriori

Estymacja maksymalnego a posteriori (3)

3. Wyznaczyć gradient zlogarytmowanego rozkładu a posteriori

$$\frac{\partial \ln p(\theta|\mathcal{D})}{\partial \theta} = \frac{a-1}{\theta} - \frac{b-1}{1-\theta} + \sum_{n=1}^{N} \frac{x_n}{\theta} - \frac{1-x_n}{1-\theta}$$

Estymacja maksymalnego a posteriori (3)

3. Wyznaczyć gradient zlogarytmowanego rozkładu a posteriori

$$\frac{\partial \ln p(\theta|\mathcal{D})}{\partial \theta} = \frac{a-1}{\theta} - \frac{b-1}{1-\theta} + \sum_{n=1}^{N} \frac{x_n}{\theta} - \frac{1-x_n}{1-\theta}$$

4. Wyznaczyć ekstremum