Maximum Likelihood Estimation

Introduction to Data Science Algorithms

Jordan Boyd-Graber and Michael Paul

SEPTEMBER 29, 2016

Continuous Distribution: Gaussian

- Recall the density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

- Taking the log makes math easier, doesn't change answer (monotonic)
- If we observe $x_1 \dots x_N$, then log likelihood is

Continuous Distribution: Gaussian

- Recall the density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1)$$

- Taking the log makes math easier, doesn't change answer (monotonic)
- If we observe $x_1 \dots x_N$, then log likelihood is

$$\ell(\mu,\sigma) \equiv -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2 \quad (2)$$

Continuous Distribution: Gaussian

- Recall the density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

- Taking the log makes math easier, doesn't change answer (monotonic)
- If we observe $x_1 \dots x_N$, then log likelihood is

$$\ell(\mu,\sigma) \equiv -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2 \tag{2}$$

Continuous Distribution: Gaussian

- Recall the density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1)$$

- Taking the log makes math easier, doesn't change answer (monotonic)
- If we observe $x_1 \dots x_N$, then log likelihood is

$$\ell(\mu,\sigma) \equiv -N \log \sigma - \frac{N}{2} \log 2\pi - \frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \quad (2)$$

Continuous Distribution: Gaussian

- Recall the density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

- Taking the log makes math easier, doesn't change answer (monotonic)
- If we observe $x_1 \dots x_N$, then log likelihood is

$$\ell(\mu,\sigma) \equiv -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2 \tag{2}$$

MLE of Gaussian μ

$$\ell(\mu,\sigma) = -N \log \sigma - \frac{N}{2} \log 2\pi - \frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \tag{3}$$

$$\frac{\partial \ell}{\partial \mu} = 0 + \frac{1}{\sigma^2} \sum_i (x_i - \mu) \tag{4}$$

MLE of Gaussian μ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2 \tag{3}$$

$$\frac{\partial \ell}{\partial \mu} = 0 + \frac{1}{\sigma^2}\sum_i (x_i-\mu) \tag{4}$$

MLE of Gaussian μ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i - \mu)^2 \tag{3}$$

$$\frac{\partial \ell}{\partial \mu} = 0 + \frac{1}{\sigma^2}\sum_i (x_i - \mu) \tag{4}$$

MLE of Gaussian μ

$$\ell(\mu, \sigma) = -N \log \sigma - \frac{N}{2} \log 2\pi - \frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \tag{3}$$

$$\frac{\partial \ell}{\partial \mu} = 0 + \frac{1}{\sigma^2} \sum_i (x_i - \mu) \tag{4}$$

Solve for μ:

$$0 = \frac{1}{\sigma^2} \sum_i (x_i - \mu) \tag{5}$$

$$0 = \sum_i x_i - N\mu \tag{6}$$

$$\mu = \frac{\sum_i x_i}{N} \tag{7}$$

MLE of Gaussian μ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2 \quad (3)$$

$$\frac{\partial \ell}{\partial \mu} = 0 + \frac{1}{\sigma^2}\sum_i (x_i-\mu) \quad (4)$$

Solve for μ:

$$0 = \frac{1}{\sigma^2}\sum_i (x_i-\mu) \quad (5)$$

$$0 = \sum_i x_i - N\mu \quad (6)$$

$$\mu = \frac{\sum_i x_i}{N} \quad (7)$$

Consistent with what we said before

MLE of Gaussian σ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i - \mu)^2 \tag{8}$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{N}{\sigma} + 0 + \frac{1}{\sigma^3}\sum_i (x_i - \mu)^2 \tag{9}$$

MLE of Gaussian σ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i - \mu)^2 \tag{8}$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{N}{\sigma} + 0 + \frac{1}{\sigma^3}\sum_i (x_i - \mu)^2 \tag{9}$$

MLE of Gaussian σ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i - \mu)^2 \tag{8}$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{N}{\sigma} + 0 + \frac{1}{\sigma^3}\sum_i (x_i - \mu)^2 \tag{9}$$

MLE of Gaussian σ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2 \tag{8}$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{N}{\sigma} + 0 + \frac{1}{\sigma^3}\sum_i (x_i-\mu)^2 \tag{9}$$

MLE of Gaussian σ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i - \mu)^2 \tag{8}$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{N}{\sigma} + 0 + \frac{1}{\sigma^3}\sum_i (x_i - \mu)^2 \tag{9}$$

Solve for σ:

$$0 = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_i (x_i - \mu)^2 \tag{10}$$

$$\frac{N}{\sigma} = \frac{1}{\sigma^3}\sum_i (x_i - \mu)^2 \tag{11}$$

$$\sigma^2 = \frac{\sum_i (x_i - \mu)^2}{N} \tag{12}$$

MLE of Gaussian σ

$$\ell(\mu,\sigma) = -N\log\sigma - \frac{N}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2 \quad (8)$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{N}{\sigma} + 0 + \frac{1}{\sigma^3}\sum_i (x_i-\mu)^2 \quad (9)$$

Solve for σ:

$$0 = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_i (x_i-\mu)^2 \quad (10)$$

$$\frac{N}{\sigma} = \frac{1}{\sigma^3}\sum_i (x_i-\mu)^2 \quad (11)$$

$$\sigma^2 = \frac{\sum_i (x_i-\mu)^2}{N} \quad (12)$$

Consistent with what we said before