

## Parameter Estimation - MLE and MAP

Parameter Estimation: Inferring  $\theta$  from observed data

### Maximum Likelihood Estimation (MLE)

#### Objective

$$\operatorname{argmax}_{\theta} L(\theta | \text{data})$$

#### Likelihood Function

$$L(\theta) = P(\text{data} | \theta) = \prod_i P(x_i | \theta)$$

#### Log-Likelihood

$$\ell(\theta) = \log L(\theta)$$

Easier to optimize

#### MLE for Normal Distribution

$$\hat{\mu} = \bar{X}$$

$$\hat{\sigma}^2 = (1/n) \sum (x_i - \bar{X})^2$$

### Maximum A Posteriori Estimation (MAP)

#### Objective

$$\operatorname{argmax}_{\theta} P(\theta | \text{data})$$

#### Using Bayes' Theorem

$$P(\theta | \text{data}) = P(\text{data} | \theta) P(\theta) / P(\text{data})$$

MAP = MLE + Prior

Incorporates prior knowledge

#### Key Difference

### Quick Comparison

| Aspect | MLE       | MAP                    |
|--------|-----------|------------------------|
| Prior  | No prior  | Uses prior $P(\theta)$ |
| Focus  | Data only | Data + knowledge       |

### Regression Use

MLE/MAP estimate regression coefficients