SESSION 20

17 April 2025 12:19

Connection of MLE and Optimization

(Boyd 351 - 357)

MLE : let us consider a family of probablity distribution on $\mathbb{R}^m$ with densities $P_x(\cdot)$

Let y be an observed sample from $P_x(\cdot)$

When consider $P_x(y)$ as a function of x for fixed $y \in \mathbb{R}^n$, $P_x(y)$ is called likelihood.

Define log- likelihood function $l(x) = \log(P_x(y))$

Consider the problem of estimating the parameter x based on observed y.

MLE estimation of $x = \hat{x}_{ml} = \arg\max P_x(y)$

Suppose we have some constraints on $x \in C \subseteq \mathbb{R}$

ML estimator of the parameter x will be

$$\max l(x) = \max \log P_x(y)$$

$$\text{subject to} \quad x \in C$$

This is an optimization problem with variable x.

The ML Problem (1) is a convex optimization problem if

(i) the log likelihood function l is convex for each value of y.

(ii) the constraint set C described by linear equality and convex

(i) the log likelihood function l is convex for each value of y.

(ii) the constraint set C described by linear equality and convex

Then the ML estimate can be computed by convex optimization.

<u>Example</u> :- Model $y_i = a_i^T x + e_i$, $i = 1 (1) m$

where $x \in \mathbb{R}^n$ is parameter, unknown quantitice

$y_i \in \mathbb{R}$ observed quantities

$a_i \in \mathbb{R}^n$ Known quantity & e_i error (IID) with density p on $\mathbb{R}$

The likelihood function $P_x(e) = \prod_{i=1}^{m} P(y_i - a_i^T x)$

ML estimator is any optimal point of the problem $\max_x \prod_{i=1}^{m} p(y_i - a_i^T x)$

Specific example suppose the the error is exponentially dist^n with the pdf.

$$p(z) = \begin{cases} \frac{1}{a} e^{-\frac{z}{a}} & z \geq 0 \\ 0 & z < 0 \end{cases}$$

Now solve ML estimator of x

from the previous discussion of constraint optimization

$$\hat{x}_{ML} = \max_x \sum_{i=1}^{m} \log P(y_i - a_i^T x_i) \quad \text{Subject to} \quad y_i - a_i^T x \geq 0$$

where $z_i = y_i - a_i^T x$

$$\hat{x}_{ML} = \max_x \sum_{i=1}^{m} \log \left(\frac{1}{\lambda} e^{-\frac{(y_i - a_i^T x)}{\lambda}} \right) \quad \text{subject to } y_i \geq a^T x \ \forall i$$

$$\equiv \max_{x} \sum_{i=1}^{m} \left\{ -\log \lambda - \frac{(y_i - a_i^T x)}{\lambda} \right\}$$ [from the exponential density of error ρ, λ known]

$$\equiv \min_{x} \sum_{i=1}^{n} (y_i - a_i^T x) \quad \text{subject to} \quad y_i \geq a^T x_i$$

$$\equiv \min_{x} \mathbb{1}^T (y - Ax) \quad \text{subject to} \quad Ax \leq y$$ $\rightarrow$ This is a LP