Johns Hopkins Engineering

625.464 Computational Statistics

The Role of Optimization in Inference

Module 2 Lecture 2A

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

The Role of Optimization in Inference

Notation:

$$Y \qquad (y_1, \ldots, y_n)$$

$$\theta \in \Theta$$

$$S(\theta) \qquad S: \Theta \to \mathbb{R}$$

$$\hat{\theta} \qquad \theta_*$$

2 Types of Optimization Problems.

1. minimization

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}}\ S(\theta)$$

2. maximization

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}}\ S(\theta)$$

Where does computational statistics intervene?

Answer: In statistical problems where s(θ) doesn't behave.

Estimation by Minimum Residuals.

$$E[y] = f(\theta_*)$$

observations: $(y_1, \dots, y_n)$

$$r_i(\theta) = y_i - f(\theta)$$

$$(y_1, \ldots, y_n)$$

minimize $r_i(\theta) = y_i - f(\theta)$ over all θ

$$\vec{r}(\theta) = \begin{bmatrix} r_1(\theta) \\ \vdots \\ r_n(\theta) \end{bmatrix}$$

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \|r(\theta)\| = \underset{\theta \in \Theta}{\operatorname{argmin}} S_p(\theta)$$

where $S_p(\theta) = \sum_{i=1}^{n} |y_i - f(\theta)|^p$

Let $\rho(\cdot)$ be some nonnegative function then minimizing

$$S_\rho(\theta) = \sum_{i=1}^{n} \rho(y_i - f(\theta))$$

M-estimator

Many functions can be optimized analytically.

Find min $f(x) = (x-1)^2$

$$f'(x) = 2(x-1)$$

$$f'(x) = 0 \quad x = 1$$

$\uparrow$ minimum

$$f''(x) = 2$$

$$g(x) = \frac{\log x}{1+x} \quad g'(x) = \frac{1 + 1/x + \log x}{(1+x)^2} \overset{?}{=} 0$$

Assumptions

1. Want to optimize $g(x)$ w.r.t. x (p-dim)
2. assume maximization
3. g is smooth and differentiable

$$g'(\vec{x}) = \vec{0}$$

1. x_0
2. done

x_0 $\quad x_t, t = 1, 2, 3, \ldots$