

Lecture Note 6: Maximum Likelihood Estimation

Least squares vs MLE

Random var. X w/ pdf $f(x; \theta)$

Sample X_i iid: $\{X_i\}_{i=1}^N \rightarrow$ each i has x_i

Likelihood:

$$\max_{\tilde{\theta}} L = \max_{\tilde{\theta}} \prod_{i=1}^N f(x_i; \tilde{\theta})$$

$$\max_{\tilde{\theta}} \ln L = \max_{\tilde{\theta}} \sum_{i=1}^N \ln[f(x_i; \tilde{\theta})]$$

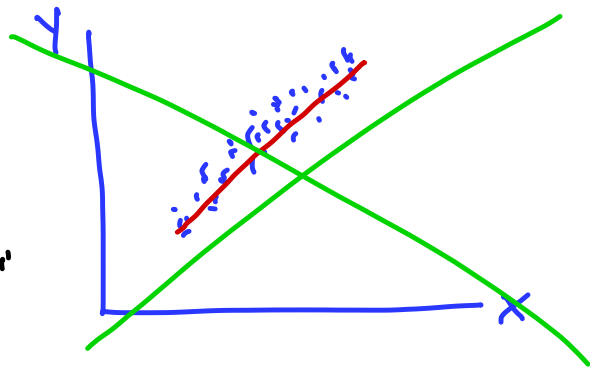
Solution: $\hat{\theta} = \tilde{\theta}^*$ satisfies $\partial \ln L / \partial \hat{\theta} = 0$

Properties of MLE: ① Consistency: $\hat{\theta} \rightarrow \theta$

② Asymptotic normality (CLT):

$$\hat{\theta} \rightarrow N(\theta, \Sigma)$$

③ Asymptotic efficiency



Bernoulli $X_i = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$

Sample $N=3$, iid, $(1, 1, 0)$

$$\begin{aligned} L &= \Pr[X_1=1] \cdot \Pr[X_2=1] \cdot \Pr[X_3=0] \\ &= p \cdot p \cdot (1-p) \\ &= p^2(1-p) \end{aligned}$$

MLE: $\max_{\tilde{p}} L = \max_{\tilde{p}} \tilde{p}^2(1-\tilde{p})$

$$\max_{\tilde{p}} \ln L = \max_{\tilde{p}} 2 \ln(\tilde{p}) + \ln(1-\tilde{p})$$

FOC: $\frac{d \ln L}{d \tilde{p}} = \frac{2}{\tilde{p}} - \frac{1}{1-\tilde{p}} = 0 \Rightarrow \hat{p} = \frac{2}{3}$

N Bernoulli variables, S successes, $N-S$ failures

$$L = p^S (1-p)^{N-S}$$

$$\ln L = S \ln(p) + (N-S) \ln(1-p)$$



Red arrows point from the L in the first equation to the $\ln L$ in the second equation, and from the $\ln L$ to the \hat{p} in the third equation.

$$\hat{p} = \frac{S}{N}$$

Methods for finding the MLE

- ① Analytic optimization
- ② (Undirected) grid search
- ③ Directed search

