

Statistics – Information Criterion

Yabusame

2024-04-25

1 信息准则的几个谣传

AIC 和 BIC 的原始形式如下:

$$\text{AIC}_m = -2 \sum_{i=1}^n \log(p_{\hat{\theta}}(y_i)) + 2d_m$$

$$\text{BIC}_m = -2 \sum_{i=1}^n \log(p_{\hat{\theta}}(y_i)) + d_m \log(n)$$

1.1 AIC 适合预测, BIC 适合解释 ?

这种看法忽略了参数化和非参数化情形, AIC 可能在 $n \rightarrow \infty$ 时也不能取到最好模型, 在非参数情况下, BIC 选择生成数据模型的一致性也不是良好定义的。

1.2 应该使用 AIC, 因为现实情况更常见到非参数化的情形 ?

We can nest subproblems!

1.3 l_0 不如 LASSO, SCAD, MCP, 因为它是离散的 ?

As far as we want!