Persado Email Subject Lines

Digital and Algorithmic Marketing (37304-01)

Brian Chingono, Will Clark, Matthew DeLio, Jonathan Stevens

CHICAGO BOOTH

The University of Chicago Booth School of Business

May 20, 2016

Modeling objectives:

- Predict click behavior out of sample
  select variables via cross validation
- Prefer simplicity/parsimony to complexity
  choose a small model
- Main effects matter more than interaction effects
  impose a bias toward main effects

The model:

$$\frac{\Pr(\text{click}=1)}{1-\Pr(\text{click}=1)} = \beta_0 + \underbrace{\sum_{j=1}^{p} x_j\beta_j}_{\text{main effects}} + \underbrace{\sum_{j=1}^{p}\sum_{k=1}^{j} x_j x_k \beta_{jk}}_{\text{interaction effects}} + \varepsilon$$

The lasso:

$$\min_{\beta}\left(-\frac{2}{n}\log \text{LHD}(\beta) + \underbrace{\lambda\sum_{i}|\beta_i|}_{\text{main + int. effects}}\right)$$

We can vary the weight on λ to impose different shrinkage for different variables!

The lasso with a twist:

$$\min_{\beta} \left(-\frac{2}{n} \log \mathrm{LHD}(\beta) + \underbrace{\omega_1 \lambda \sum_{j=1}^{p} |\beta_j|}_{\text{main effects}} + \underbrace{\omega_2 \lambda \sum_{j=1}^{p} \sum_{k=1}^{j} |\beta_{jk}|}_{\text{interaction effects}} \right)$$

$\omega_1 < \omega_2$: the lasso will penalize interaction effect parameters more than main effect parameters (*impose a bias toward main effects*)

Trade-off between accuracy (*select variables via cross validation*) and simplicity/parsimony (*choose a small model*)