Decision Trees II

Model-Based Recursive Partitioning

Introduction

Extending Decision Trees

- Conditional Inference Trees (Hothorn et al. 2006)
  - Addresses selection bias for variables with many potential split points
  - Separates variable and split point decision
  - Variable selection and stopping criterion based on statistical test
- Model-based Recursive Partitioning (Zeileis et al. 2008)
  - Connects recursive partitioning with fitting parametric (regression) models
  - Approach to fitting "homogeneous" models in tree nodes

Model-based Recursive Partitioning

Model-based Recursive Partitioning

- Specify a parametric model in advance
  - Constant leaves: $y \sim \mathcal{N}(\theta, \sigma^2)$
  - OLS regression: $y|x \sim \mathcal{N}(x'\theta, \sigma^2)$
  - logit, survival, ...
- Three types of variables
  - Response variable y
  - Predictor variables $x_1, \ldots, x_p$
  - Partitioning variables $z_1, \ldots, z_j$

$\rightarrow$ Use the partitioning variables $z_1, \ldots, z_j$ to partition the n observations into distinct groups with **separate models**

Model-based Recursive Partitioning

- Specify a parametric model in advance
  - Constant leaves: $y \sim \mathcal{N}(\theta, \sigma^2)$
  - OLS regression: $y|x \sim \mathcal{N}(x'\theta, \sigma^2)$
  - logit, survival, ...
- Three types of variables
  - Response variable y
  - Predictor variables $x_1, \ldots, x_p$
  - Partitioning variables $z_1, \ldots, z_j$

$\rightarrow$ Use the partitioning variables $z_1, \ldots, z_j$ to partition the n observations into distinct groups with **separate models**

Model-based Recursive Partitioning

Algorithm 1: Recursive partitioning with GLMs

Parameter : p-value threshold
Initialization: Fit initial model using all observations

Perform M-fluctuation tests for each partitioning variable (H_0: stability of $\hat{\beta}$);
if *minimum p-value exceeds threshold* **then**
    end partitioning (global H_0 not rejected);
else
    choose partitioning variable associated with the smallest p-value;
    find the optimal split point;
    split node into two subnodes at this split point;
    for *each node of the current tree* **do**
        continue partitioning process;
    end
end

Model-based Recursive Partitioning

M-fluctuation tests

Starting point: Objective function of OLS or ML

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{n} \Psi(y_i, x_i, \theta)$$

Solved via score function

$$\frac{\partial \Psi(y_i, x_i, \theta)}{\partial \theta}$$

Partial derivations (e.g., $y = \beta_0 + \beta_1 x + \epsilon$)

$$\frac{d\Psi(y_i, x_i, \beta_0, \beta_1)}{d\beta_0} = \epsilon_i \quad \text{and} \quad \frac{d\Psi(y_i, x_i, \beta_0, \beta_1)}{d\beta_1} = x_i \epsilon_i$$

$\rightarrow$ Order contributions to score function evaluated at the current estimates with respect to partitioning variable to detect **structural change**

Model-based Recursive Partitioning

M-fluctuation tests

Starting point: Objective function of OLS or ML

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \sum_{i=1}^{n} \Psi(y_i, x_i, \theta)$$

Solved via score function

$$\frac{\partial \Psi(y_i, x_i, \theta)}{\partial \theta}$$

Partial derivations (e.g., $y = \beta_0 + \beta_1 x + \epsilon$)

$$\frac{d\Psi(y_i, x_i, \beta_0, \beta_1)}{d\beta_0} = \epsilon_i \quad \text{and} \quad \frac{d\Psi(y_i, x_i, \beta_0, \beta_1)}{d\beta_1} = x_i \epsilon_i$$

→ Order contributions to score function evaluated at the current estimates with respect to partitioning variable to detect **structural change**

Model-based Recursive Partitioning

M-fluctuation tests

Starting point: Objective function of OLS or ML

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{n} \Psi(y_i, x_i, \theta)$$

Solved via score function

$$\frac{\partial \Psi(y_i, x_i, \theta)}{\partial \theta}$$

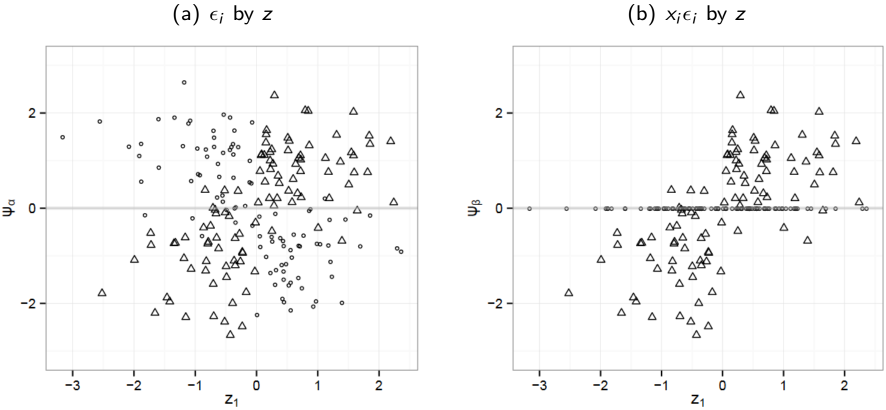
Partial derivations (e.g., $y = \beta_0 + \beta_1 x + \epsilon$)

$$\frac{d\Psi(y_i, x_i, \beta_0, \beta_1)}{d\beta_0} = \epsilon_i \quad \text{and} \quad \frac{d\Psi(y_i, x_i, \beta_0, \beta_1)}{d\beta_1} = x_i \epsilon_i$$

$\rightarrow$ Order contributions to score function evaluated at the current estimates with respect to partitioning variable to detect **structural change**

Model-based Recursive Partitioning

Figure: Structural change illustration with x_i\{treatment (Δ), control ($\circ$)\}



Seibold et al. (2016)

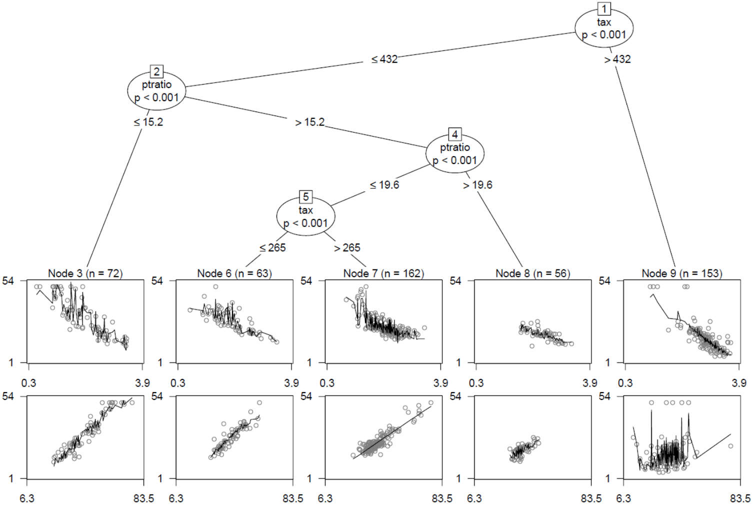
Model-based Recursive Partitioning

Controlling tree complexity

- Test-based stopping
  - α (`alpha`, default $= 0.05$)
- Tree-based stopping
  - Observations per node (`minsize`)
  - Tree depth (`maxdepth`)
- Post-pruning
  - Prune nodes by checking AIC or BIC improvements

Model-based Recursive Partitioning

Figure: Linear-regression-based tree with partial scatterplots



Zeileis et al. (2008)

Summary

- Recursive partitioning approach can be complemented with permutation and M-fluctuation tests
- Adjusts and (heavily) extends CART idea
- Model-based Recursive Partitioning
  1. Tree-based approach to explore interactions
  2. Can be used to detect model misspecification (Kopf et al. 2010)
  3. Extensions for various types of models available (e.g. Strobl et al. 2015)

Software Resources

Resources for R

- A Laboratory for Recursive Partytioning: `party`
  - Includes `ctree` (Conditional Inference Trees) and `mob` (Model-based Recursive Partitioning)
- A Toolkit for Recursive Partytioning: `partykit`
  - Includes a re-implementation of `ctree` and new interfaces for `mob` (`lmtree`, `glmtree`)

References

Hothorn, T., Hornik, K., Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics 15*(3), 651–674.

Kopf, J., Augustin, T., and Strobl, C. (2010). *The Potential of Model-based Recursive Partitioning in the Social Sciences – Revisiting Ockham's Razor*. Technical Report Number 88, University of Munich.

Seibold, H., Zeileis, A., Hothorn, T. (2016). Model-Based Recursive Partitioning for Subgroup Analyses. *The International Journal of Biostatistics 12*(1), 45–63.

Strobl, C., Kopf, J., Zeileis, A. (2015). Rasch Trees: A New Method for Detecting Differential Item Functioning in the Rasch Model. *Psychometrika 80*(2), 289–316.

Zeileis, A., Hothorn, T., Hornik, K. (2008). Model-Based Recursive Partitioning. *Journal of Computational and Graphical Statistics 17*(2), 492–514.

Zeileis, A., Hornik, K. (2007). Generalized M-fluctuation tests for parameter instability. *Statistica Neerlandica 61*(4), 488–508.