

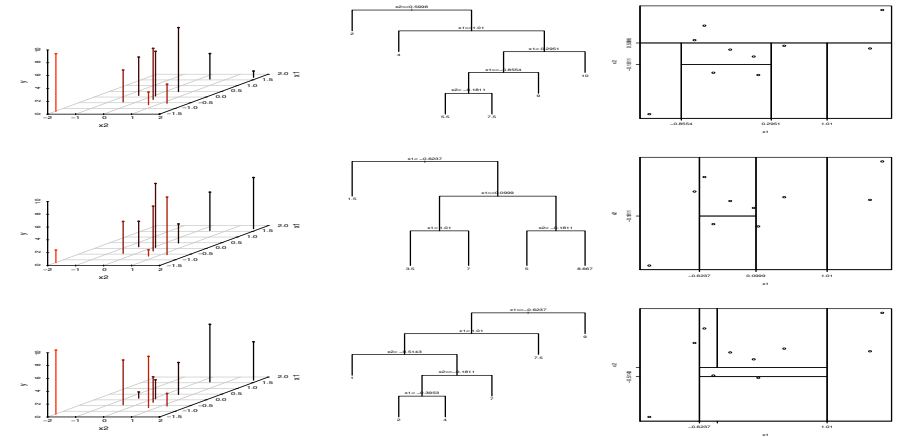
## Matching and ATT Estimation via Random Recursive Partitioning

Wien April 2006

RRP is non standard application of CART: we grow CART's with

```
> y <- runif(NROW(mydata))
```

```
> rpart(y ~ ., mydata)
```

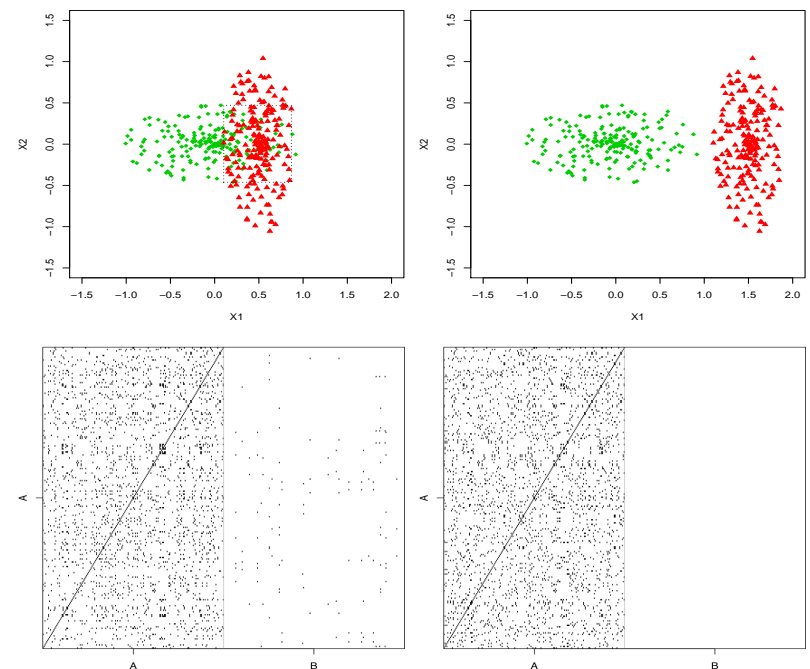


## What is RRP for?

At each random partition we set  $\pi_{ij} = 1$  if obs.  $i$  and  $j$  are in the same cell. At the end we average out.

This way we generate a proximity matrix (but this is not Leo Breiman's proximity in RandomForest) that can be used in:

- analysis of the common support between dataset (how much do they overlap?)
- matching and ATT estimation
- supervised and unsupervised classification
- (hot-deck) missing data imputation

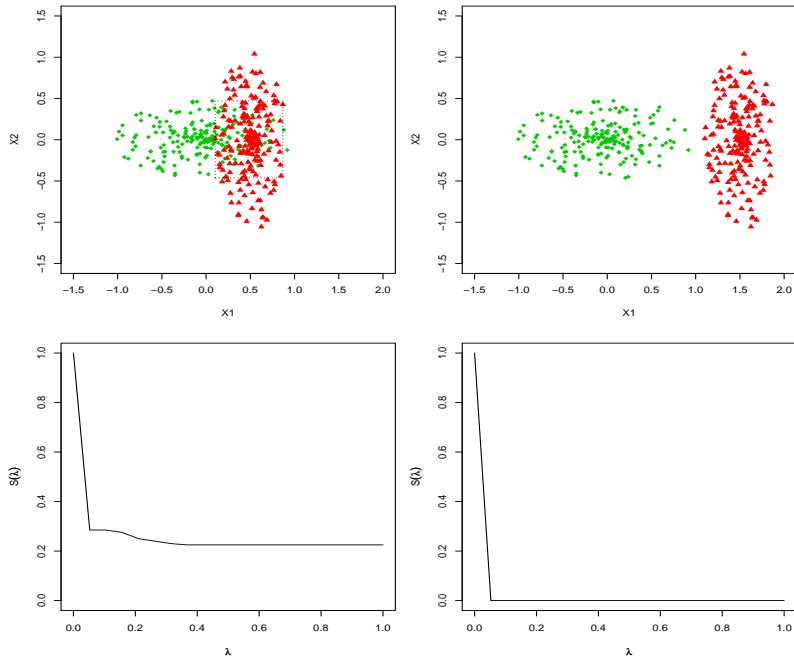


# The RRP package

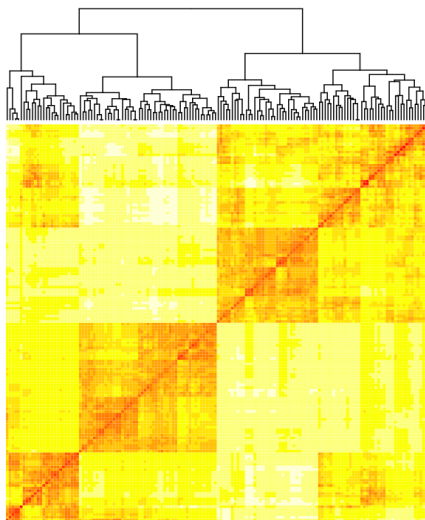
<code>rrp.dist</code>	dissimilarity matrix, obj of class 'dist'
<code>rrp.class</code>	nearest neighbor classifier
<code>rrp.predict</code>	nearest neighbor predictor
<code>rrp.impute</code>	hot-deck nearest neighbor imputation

Related papers at

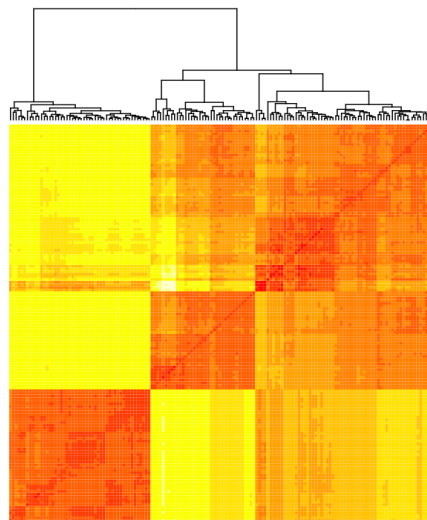
<http://services.bepress.com/unimi>



## RRP vs Euclidean clustering on Iris data (average link)



RRP  
7 miss-classified



Euclidean  
14 miss-classified