Remote Sensing and Image Analysis – Supervised classification methods

Edzer Pebesma

November 26, 2008

ifgi
Institute for Geoinformatics
University of Münster

Supervised classification – what is it?

Discriminant analysis, “Maximum likelihood classification”, k-NN, support vector machines, ...

- who has used it?
- what for?
- what does it do?
- what does it *not* do?

What does it do?

- Goal: classification, i.e. prediction of a categorical variable
- Data reduction: n continuous variables to 1 p-class categorical
- usually/often ignores spatial organization of data
- needs observations on the class variable ("ground truth", "control points")

It's place in multivariate statistics

- supervised: we distinguish between dependent (class) and independent (bands)
- classify: from continuous variability, go to classes
  - the question is not: "are the data grouped" (cluster analysis)
  - the question is: can we predict this particular categorical variable, based on (usually continuous) variables (bands).

Discriminant analysis

- result: discriminant axes ($p-1$, with p number of classes)
- plot of data in LD1/LD2, LD1/LD3, LD2/LD3, ...
- prediction: for *any* point in feature (spectral) space
  - class (with maximum a posteriori probability)
  - probability of belonging to each of the classes
  - ... ambiguity of choice
- valuation/error analysis by cross validation; for each observation:
  - leave observation out
  - compute LD's
  - compare predicted class with observed class
- LDA and QDA: LDA uses a single covariance matrix (equal distribution per category), QDA uses a covariance matrix for each group.

k-NN

- for each point in feature (spectral) space, find the k nearest neighbours
- classify according to majority in this selection.
- local, non-parametric
- k is tuning parameter
- needs distance measure in feature space: do scaling?
- does not address correlation among features: do PCA first?
- simple, intuitive
- algorithm: needs (k-D?) tree indexes for high-dimensions and massive data sets

RandomForest

- generalizes k-NN
- builds many trees on random (resampled) permutations of the data
- forms a classifier that merges all trees, weighted by how well each of them fitted the data
- function `randomForest` in R package `randomForest`

Classification tree

Classification and Regression Trees (CART)

- cuts the feature space by consecutively cutting a feature variable
- simple, explainable/graphable, attractive.
- ignores joint variability; essentially stepwise uni-variate!

Support vector machine (svm)

New, smart, price-winning

- finds the support vector, that best splits two groups either:
  - maximize distance between boundary points
  - penalize mis-classifications
- uses linear vectors but in (polynomial) manyfolds of the feature space; this means possibly very complex separating planes
- can do regression too!
- fast!
- function `svm` in R package `e1071`