LECTURE 2 2: SUMMARY

STAT 598z: Introduction to computing for statistics

Vinayak Rao

Department of Statistics, Purdue University

April 27, 2017

Class overview

We looked at:

- Basics of R programming
  - Data structures
  - Control structures
  - Functions

Class overview

We looked at:

- Basics of `R` programming
  - Data structures
  - Control structures
  - Functions
- Slightly more advanced topics
  - plotting with `ggplot`
  - Object oriented programming
  - regular expressions
  - Data manipulation with tidyverse

Class overview

We looked at:

- Basics of `R` programming
  - Data structures
  - Control structures
  - Functions
- Slightly more advanced topics
  - plotting with `ggplot`
  - Object oriented programming
  - regular expressions
  - Data manipulation with tidyverse
- Some computational ideas
  - Crossvalidation
  - Ridge regression and LASSO
  - Monte Carlo

Topics we did not cover

- Interfacing with C
- Parallel computing
- Writing your own packages

Topics we did not cover

- Interfacing with C
- Parallel computing
- Writing your own packages

Hopefully you're confident enough to read about this yourself

Topics we did not cover

- Interfacing with C
- Parallel computing
- Writing your own packages

Hopefully you're confident enough to read about this yourself

Also, confident to teach yourself other languages

- python
- Matlab
- scala/julia/other more exotic languages
- C

A more algorithmic course on computational statistics

- Optimizing parameters with missing variables (EM algorithm)
- Dealing with hidden variables for structured problems
  - Baum-Welch algorithm for hidden Markov models
  - Kalman-filtering
- More Monte Carlo and MCMC, especially in the context of Bayesian computation
- More optimization algorithms
- Loss functions beyond LASSO

Projects

Submit a 4-8 page report detailing

- What you wanted to do
- What you did
- Some results and nice plots
- What you couldn't do (and why)
- What you learnt

Don't include code in your report (submit separately)