Statistical Computing 2: Course Overview

Matteo Fasiolo

University of Bristol

Overview

During this course you will learn to:

1. interface R with C++ to perform computational intensive tasks
2. perform parallel computations by
   - using the OpenMP or Intel TBB API in your code
   - launching parallel jobs on a cluster (Bluecrystal)
3. program in Python:
   - from basic data analysis to deep learning

Overview

The detailed programme is (week by week):

- ⑬ Intro to C++
    - **Thu** lab on C++ (ACRC)
- ⑭ Intro to R/C++ integration using Rcpp package (Matteo)
    - **Tue** Lecture on basic Rcpp
    - **Thu** Lab on basic Rcpp
- ⑮ Simulations and linear algebra in Rcpp (Matteo)
    - **Tue** Lecture on Rcpp Sugar and RcppArmadillo
    - **Thu** Lab on advanced Rcpp

Note aim is **interfacing** C++ with R, not building stand-alone C++ programs.

Overview

16. Intro to Linux
    - **Tue** using C++ in R packages (Matteo)
    - **Thu** Intro to Linux (ACRC)
17. Intro to HPC
    - **Tue** Office Hours (Matteo)
    - **Thu** intro to HPC using Blue Crystal (ACRC)
18. Intro to OpenMP
    - **Tue** Office Hours (Matteo)
    - **Thu** intro to parallelization using OpenMP (ACRC)
19. Intro to Intel-TBB
    - **Tue** Office Hours (Matteo)
    - **Thu** intro to parallelization uing Intel-TBB (ACRC)

Matt's labs are in TYNDALLS PARK RD 30-32 2.03 2-5pm.

Overview

19. Introduction to Deep Learning
    - **Tue** Office Hours (Matteo)
    - **Thu** Introduction to Deep Learning (ACRC)
20. Data Analysis in Python (SCRAPPED)
    - **Tue** Office Hours (Matteo)
    - **Thu** Data Analysis in Python (ACRC)
21. Parallelization in Rcpp
    - **Tue** Intro to RcppParallel (Matteo)
    - **Thu** Lab on RcppParallel (Matteo)

Weeks 19 and 20 dedicated to project work.

Learning material

Material on interfacing R with C++ is here:

https://mfasiolo.github.io/sc2-2019/

You can submit pull requests here:

https://github.com/mfasiolo/sc2-2019

You should be able to download the repository and run the code.

You can also get the ebook:
"Seamless R and C++ Integration with Rcpp" (Dirk Eddelbuettel)
from the university library website.

The material for the ACRC labs is here:

http://www.bris.ac.uk/acrc/acrc-training/
https://chryswoods.com/main/courses.html

Assessment

Assessment is the same as for SC1:

- personal portfolio 30%
  - send it to matteo.fasiolo@bristol.ac.uk
  - at least 2 pages PDF per topic (Rmarkdown or Jupyter Notebook)
  - can include accompanying C++ code
- assessed coursework 40%
  - send it to matteo.fasiolo@bristol.ac.uk
  - choose two topics to explore in depth
  - longer PDF document
  - if R-based, submit an accompanying R package*
  - if Python-based, a Jupiter Notebook document or Python package*
- group work 30%
  - submitted via Blackboard
  - joint between SM2 and SC2
  - either an R-package or a Python package on GitHub

Deadline is June 5th.

* Verify system requirements, they must work on other people's computer!

Assessment

Last three years most group projects used Rcpp!

It's a rather useful package:

https://cran.r-project.org/web/packages/Rcpp/index.html

Note that these are only direct dependencies.