Quantitative Trading With R

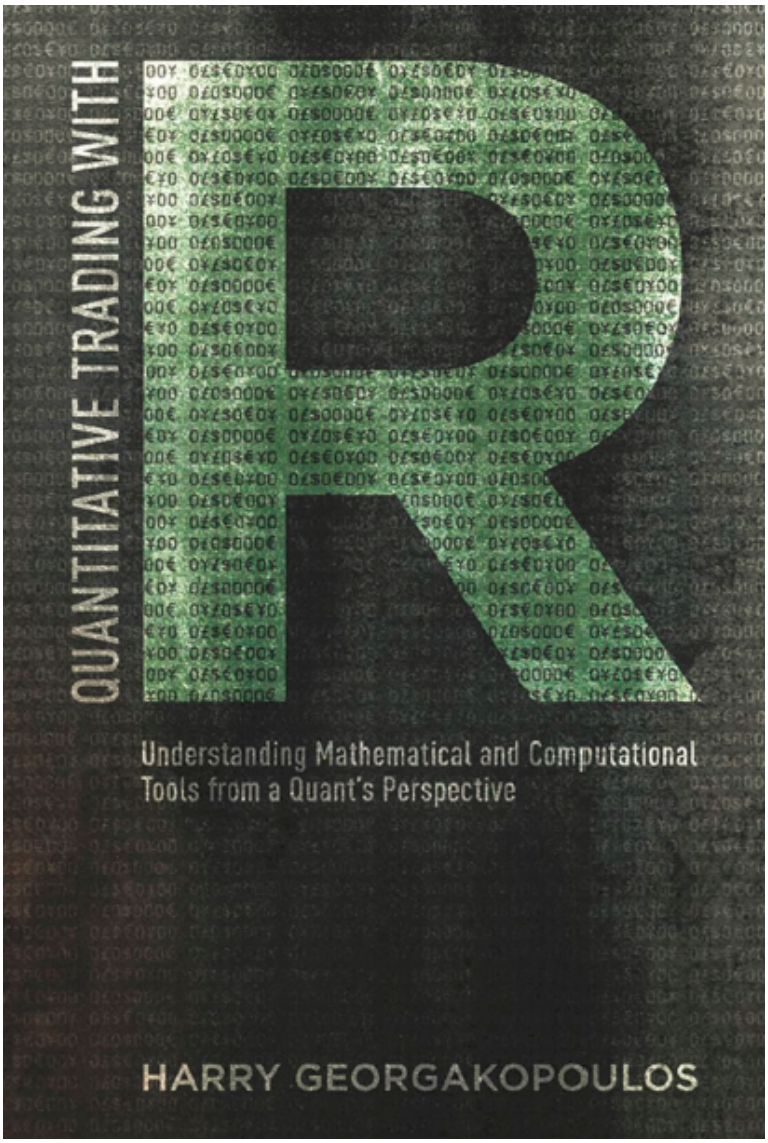
Summary

RK

April 12, 2015

Abstract

The purpose of this document is to summarize the book written by Harry Georgakopoulos, titled, "Quantitative Trading With R".



Summary

Chapter 1 provides an overview of the book in which the author clearly mentions that this book should be considered as an introduction to quantitative trading. One thing that strikes me odd; Why should you price a book that is supposedly an introduction to R and quant trading at $ 66.50. This was the price that I saw when I last checked out on Amazon. May be the author has absolutely no say in the pricing of the book. Pirated copies of the book are already available on the net. I have read through the book and I think that the author should have been more honest about the contents of the book. May be he should have called the book *A loose introduction to A serious introduction of Quant trading with R*

The book is close to 240 pages long and the first 120 pages of the book is on R. This is absolutely crazy. One could have pushed all this to appendix spanning no more than 10 pages. Using 50% of the book to educate the readers about R is totally useless, given the grand title of the book. A person who is absolutely new to R will hardly learn anything from 120 pages. I think author should make it absolutely clear that a basic working knowledge of R is needed to start with the book and use the space available to focus on the subject that he is dealing with. I have seen many authors using this tactic of producing books with titles such as, "Learning blah-blah using R". These authors take up enormous space focusing on the latter part and rush through the former part. The publishers for some reason think that they can charge insane prices. This pattern will surely disappear as more and more readers realize that they are being tricked by such authors and publishers.

Chapter 2 covers `base R` in 30 pages, pitifully inadequate.

Chapter 3 covers data munging in 30 pages.The following are the packages used by the author in the garb of *educating* readers about working with data :

- `quantmod`
- `RJSONIO`
- `XLConnect`
- `RODBC`
- `RMySQL`
- `plyr`
- `xts`
- `TTR`
- `ggplot2`

I find this chapter, totally useless as it illustrate ideas at 100000 . . . ft level using tons of packages. At such extremely high altitude, the reader doesn't get to see anything worthwhile, if at all he manages to strain his eyes and looks for something.

Chapter 4 covers frequentist statistics and Bayesian statistics in 20 pages! Absolutely ridiculous. To understand the intuition behind Bayes and to get a working knowledge in Bayesian stats requires a lot of effort. The chapter trivializes the subject.

Chapter 5 covers hastily random processes in 30 pages!. The author briefly touches upon timeseries and some R functions from `urca`, `xts` and `quantmod` packages.

The first five chapters supposedly equip the reader to understand R code that is given in the later part of the book. If you already know R, the first 120 pages can be skipped completely. If you haven't coded in R before, in all probability, the first 120 pages will overwhelm you. I did not learn anything new in the first 120 pages

of the book.

Chapter 6 takes the reader through an exercise in coding a pairs trading strategy. Usually the cointegrating beta is found by a least squares regression. The chapter makes use orthogonal regression, also called as *Total Least Squares regression* to compute the hedge ratio. Out-of-sample performance for this simple strategy is explored via suitable R code.

Chapter 7 is the only chapter in the entire book that was useful to me. It is a crash course on `quantstrat` package. However most of the code is copied from the blog QuantStrat TradeR

Chapter 8 is about exploring some basic summary statistics of high frequency data. The author does not provide a link to the data used in the chapter. Hence the reader is reduced to a mere spectator in this chapter. Plots of bid price, inter arrival times between quotes, bid-ask spread, autocorrelation of micro price are shown for the dataset used in the chapter. There is also a brief mention about Lee and Ready algorithm and its improved version for inferring the trade direction. The chapter concludes by giving a reference to `highfrequency` package.

Chapter 9 explores functions from `RQuantLib` package. The functions in the package are vectorized and hence give out greeks for the options based on the type of vectorized input. Implied volatility for American options is not as straightforward as that of an European option. The package has a built-in function that gives the implied vol for American options.

Chapter 10 talks about the powerful `optim` function in `base R`. It starts off by explaining the code behind Newton Raphson method. Subsequently brute force methods are used to compute the parameters of a linear model.This is followed by a thorough explanation of `optim` function. I seriously do not know why the author has spent so much time on `optim` when R help does an amazing job of it already. The last section of the chapter illustrates quadratic programming by using functions from `DEOptim` package. A comparison between cap weighted index returns and an alternate index based on avoiding max drawdown is done to explain the usage of `DEOptim`.

Chapter 11 talks about three packages, `Rcpp`, `testthat` and `knitr`. I am happy to find that the author has mentioned `knitr` package.There is dearth of Reproducible Research documents in the trading community. In the years to come, may be the situation will improve.

Takeaway

The book is definitely overpriced for what it delivers. Given that 50% of this book explains R basics, the title is not at all appropriate. Even the quant stuff that is covered in the remaining 50% of the book is laughably inadequate.