

DDS Algorithm

Dominic Scruton & Philip Hindle

22 May 2020

Portfolio Theory

Introduction

Capital Asset Pricing Model (CAPM)

Sharpe Ratio as a Measure of Performance

The Expected return of a portfolio with k assets is simply the weighted sum of expected returns for each asset and is given by:

$$E[R_p] = \sum_{i=1}^k w_i r_i = \mathbf{w}^T \mathbf{r}$$

where \mathbf{w} is a vector of stock weights and \mathbf{r} is a vector of the expected returns for each stock. Defining these functions in matrix algebra notation emphasizes the use of vectorization within the DDS Algorithm. The variance of the expected return is then:

$$\begin{aligned} Var[E[R_p]] &= Var[\mathbf{w}^T \mathbf{r}] \\ &= \mathbf{w}^T Var[\mathbf{r}] \mathbf{w} \\ &= \mathbf{w}^T \Sigma \mathbf{w} \end{aligned}$$

Generally use simple returns when we have many assets. The portfolios Sharpe Ratio is then given as:

$$Sharpe = \frac{E[R_p] - r_f}{SD[R_p]}$$

The Sharpe ratio specifies the Additional expected return for a given level of risk. A higher Sharpe Ratio suggest a more optimal portfolio.

Methodology

Optimization Problem

Optimization is a common theme that underpins the fitting of Statistical models. From the maximization of the log-likelihood function under Linear Models, to the minimization of errors under Neural Network

classifiers, optimization is fundamentally important to obtaining objective solutions to many problems in the world of Data Science and Statistical research.

The general Optimization problem can be defined as follows:

$$x^* = \underset{x \in \Omega}{\operatorname{argmin}} f(x; y)$$

- discuss the problems- sometimes can't differentiate, other times it is not computationally feasible to find exact answers, when we have large amounts of data and lots of stocks.

In this case, one would like to maximize the Sharpe ratio such that the sum of weights in the portfolio equals one. This optimization problem can be expressed as follows:

$$\operatorname{Max} \left(\frac{E[r_p] - r_f}{\operatorname{Std}[E[r_p]]} \right) \text{ s.t. } \sum_{i=1}^K w_i = 1$$

Or in vector notation:

$$\operatorname{Max} \left(\frac{\mathbf{w}^T \mathbf{r}}{\sqrt{\mathbf{w}^T \mathbf{\Sigma} \mathbf{w}}} \right) \text{ s.t. } \|\mathbf{w}\| = 1$$

Theoretical Solution (i.e. consider the actual mathematical result)

Discretized Dynamical System (Deterministic Grid Search)

- One of the advantages is that we can set the level of granularity here. i.e. a weight of 12.4 is effectively the same as 12.378383838, as the qualitative results are effectively the same and in reality, one tends to . Really inefficient- needlessly searches the entire area and not at all granular- becomes extremely computationally inefficient for large numbers of stocks.

Bisection Method

- Identify an interval (a, b), where the gradient switches sign. Bisection is an example of a more general deterministic method

Gauss-Newton Method

Stochastic Grid Search

Simulated Annealing

The problem we face with simulated annealing, in comparison to its standard applications, is that the weights are bounded to lie in the interval $[0, 1]$ and must sum to one. Therefore, we need to decide whether to use normal distributions and truncate any weights to lie in $[0, 1]$ or use a distribution with support $[0, 1]$ (e.g. Beta distribution).

Other Optimization Methods

There exist a plethora of possible methods to solve the optimization problem identified in this report. The two general contrasting methods discussed are Stochastic and Deterministic forms of optimization.

Discuss the Advantages and disadvantages of the two methods- Deterministic and stochastic optimization methodology and also of the methods in general.

The DDS Algorithm (R Package)

Pre-Processing Function

DDS Algorithm

- Discuss how the expected returns and then the portfolio standard deviation are ANNUALIZED- see Python for Finance course

Heat Maps and Plots

Testing

To build robust software one needs to approach the problem of testing rigorously.

Unit Testing

Unit testing helps to create robust code. This ensures the code is robust to new versions of R, as well as enabling users to utilize the function correctly. This ensures the function generates the expected output and can deal with arguments of different structure, flagging problems if the structure of input arguments is not appropriate (boundary testing). This ensure code is robust and results are reproducible.

Unit testing tests functional units. For each created function, one should create further code testing whether the function works as intended. For unit testing, we typically start with function arguments such that we know the answer and then test whether the function returns the correct answer (Mailund, 2017). In this case, the problem is more complex and it is generally not possible (or not worth the time) to manually calculate the Sharpe ratio for a given set of stocks and weights. (Could either work backwards from a given Sharpe Ratio or use validated results online)

One can automate the process of testing functions by creating functions that return error messages if the results are not correct. Doing this can save time, with testing functions called whenever changes are made to each function.

In this case, because the function is not deterministic, there will be variation in the results of each function, hence a level of tolerance is set. One can also use random numbers in tests, simulating artificial data. This can often be appropriate, enabling one to test functions with more 'unusual' data, to ensure functions remain valid for outlying and extreme parameter values. If the function is not deterministic (which in our case it isn't- we have stochastic functions, all of which are approximations to the specific value).

List an example of the main data and functions used to test each function.

Boundary Testing

Other potential pitfalls of programming in R include Scoping. In R Scoping is *lexical*, so if a variable is used in a function but not defined either as an arguments or part of the function, R will start searching outward

of the function for values in the global environment (Mailund, 2017) (unlike compiled languages, such as C++ and JAVA). Therefore, incorrect function arguments or values might be called that have already been created in the global environment. To avoid this, default arguments, warning statements and stop functions are used to ensure a robust function that is not prone to user error.

Boundary tests are useful in dealing with this problem. Provide a list/table of the boundary tests used on the different functions.

Table 1- Example Boundary Tests

| Test | Function |
|---|---------------------|
| Check Arguments of correct data structure | ALL |
| Check data specified as daily returns | DDS |
| Check Starting Weights Sum to 1 | Simulated Annealing |

Simulation

- So far we have tested the function on 3 stocks, so we can conclude that syntactically there are no major issues. However, to assess the performance of the different simulation methods, one really needs to run the functions with a far larger number of stocks. As the number of stocks gets larger, ‘edges’ of the search space are less likely to be searched for random search. Furthermore, the optimization function can get stuck in local maxima, which can be problematic.
- Need to be able to simulate data such that we start with the Sharpe ratio and work backwards, to ensure that we know the exact value of the maximum Sharpe ratio and corresponding weights, in order to assess the performance of the different optimization methods.

Assessing Computational Efficiency

- discuss the computational efficiency of the different optimization methods- measure using the `system.time` function, for example.
- Also discuss how I improved the computational speed of the function using Parallelization and profiling.

The process of vectorization can be used

- Include a table for simulation speeds and results- also discuss computational complexity and efficiency

Example Portfolio

We test and simulate using stock returns from 1 January 2010 to 1 December 2020 inclusive, for Delta Airlines, Walmart and Microsoft stock in a portfolio (could add many more)

Conclusion

- Discuss the wider applicability of these methods- in Machine Learning, when we have many hyperparameters to tune, rather than using grid search, we can use Gauss- Newton and Simulated Annealing to massively reduce computationally time and search a far greater space of hyperparameter values- huge applicability in all fields of statistics and machine learning.
- Completed all of “Beginning Data Science in R”, except for the creation of packages (Chapter 11).

Bibliography

Alison Etheridge. 2002. “A Course in Financial Calculus”. Cambridge University Press.

Jonathan Berk, Peter DeMarzo. 2014. “Corporate Finance”. Pearson.

Geoffrey Grimmett, David Stirzaker. 2001. “Probability and Random Processes”. Oxford University Press.

Thomas Mailund. 2017. “Beginning Data Science in R: Data Analysis, Visualization and Modelling for the Data Scientist”. Apress.

Geof Givens, Jennifer Hoeting. 2013. “Computational Statistics”. Wiley and Sons.

Paulo Cortez. 2014. “Modern Optimization with R”. Springer.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.

Appendix

Python Code

- Attach the equivalent Python code to the above problem, however, no ‘package’ has been created- BUT could look into this.

R code

- add any extra pieces of R code here