MARKETING MATERIAL

ON THE STABILITY OF THE GENETIC ALGORITHM, APPLICATION TO US EQUITIES

Tony Guida* Philippe Huber* Federico Rossi*

June 2022

*RAM Active Investments SA, Rue du Rhône 8, CH-1204 Geneva, Switzerland

Introduction

In traditional finance, optimising a portfolio is usually a straightforward process performed with the classical Mean-Variance portfolio selection model of Markowitz. In short, it consists of taking advantage of the objective function convexity by inverting a covariance or correlation matrix to achieve an optimised portfolio. There are issues with this process that have been thoroughly studied in many articles. We will focus on the impossibility of computing a usable correlation matrix when the portfolio contains many securities, making the optimisation of large portfolios intractable by classic optimisers. Because our process generates more than 10^6 time series to build a portfolio, we turn to a Genetic Algorithm ("GA"); a technique that is widely used in sciences such as physics or medicine.

Genetic Algorithms, benefiting from their heuristic nature and inspired by the process of natural evolution, can solve portfolio construction problems that traditional methods struggle to address. They use operators such as crossover, selection and mutation to search for an optimised solution. Evolutionary algorithms, in general, require no fitness gradient information or correlation matrices to proceed, are easy to process in parallel and can escape from local minima, where deterministic optimisation methods may fail.

The aim of this paper is to focus on the stability of the Genetic Algorithm that RAM's Systematic Macro team utilises to model financial markets. Because of its randomness, one should expect to see noise in the optimisation results, especially in a very high-dimensional setup like ours, but we show that its stability is impressive when looking at an out-of-sample performance of the optimal portfolio.

Markets as networks of agents and Genetic Algorithm

Most quantitative systematic strategies follow a bottom-up technique where bets are taken on trends or reversals of securities alone, after estimating an expected alpha. Statistical arbitrage strategies will select pairs of securities that have a high correlation to each other and play the reversion of any abnormal discrepancy.

Our philosophy is different and relies on a top-down view that aims to model markets as a network. This is performed through the introduction of "agents" each of which is a long-short portfolio. Each agent takes positions based on a simple signal (e.g. closing price) in a sub-universe that is a subset of the total universe. It is categorised as either trend following or mean reverting and can be short-, mid-, or long-term.

To simulate the markets network, a Graphics Processing Unit (GPU) proprietary process generates billions of agents to cover all possible combinations of signals, sub-universes, and frequencies. A few million agents are then pre-selected depending on their in-sample behaviour and finally, our proprietary Enhanced Genetic Algorithm builds an optimised portfolio of agents.

A Genetic Algorithm is a heuristic optimisation process that relies on Darwin's theory of evolution. It uses a space search to find a near-optimal solution by running a random procedure with genetic operators like crossover, mutation, and selection[1]. It is particularly well-suited here because it deals with any number of portfolio components, and it runs very well on a GPU. Note that the Genetic Algorithm we are using in this paper has been enhanced to deal with the over-fitting risk of having such a massive number of agents.

Data and Simulation

Our universe contains 100 US single-name equities that have the highest market capitalisation. All time series' start in 1988 with a daily frequency. As explained earlier, to model the market network, we simulate agents which are defined by a sub-universe, a signal, and a time frequency. In the present study, we focus on a single signal which is the closing price. Frequencies (or, more accurately, look-back periods) go from 2 days to 1000 days in the past. 150 random sub-universes are generated with random sizes between 30 and 100 securities.

The stability of our GA is tested along two directions: among a single set of 150 sub-universes and across 6 different sets of 150 sub-universes each. For each set, we run 100 simulations. To produce a meaningful simulation, the process is run on a yearly rolling basis so that it can adapt over time: we use the data from 2000 to 2009 to simulate a portfolio in 2010, then data from 2000 to 2010 to simulate 2011, and so on, until 2021.

Trading costs and slippage are included in the results.

Results

Each run of the GA algorithm produces a daily portfolio of securities from January 2010 to December 2021. We focus on the following metrics to quantify the algorithm's stability:

[1] For more details on Genetic Algorithms, refer to our 2019 paper "Genetic Algorithms: A Heuristic Approach To Multi-Dimensional Problems" at: `https://www.ram-ai.com/sites/default/files/2020-04/201908_genetic-algorithms-academic-paper.pdf`

annualised rate of return, annualised volatility, risk-reward, percentage of up and down days, the average rate of return of up and down days, and worst drawdown.

Table 1 displays statistics from the metrics of the first set of sub-universes (SU1). It shows low values of standard deviations which is what we hoped for regarding stability. First and third quartiles further evidence the tightness of the metrics' range.

We follow the same procedure for 5 more sets of sub-universes ($SU2$ to $SU6$) and compare their metrics (Tables 2 and 3). As for the first dataset, they globally display a small standard deviation confirming the stability of the Genetic Algorithm. Moreover, results of the 6 datasets are close to each other as Figures 1 to 4 illustrate. Historical performances and correlations display the process' stability even further. It could be surprising to the reader because agents have been generated on different random sub-universes. We argue that it demonstrates how efficient our whole process is to extract information from a market seen as a network. Having an optimisation algorithm that is based on randomness, added to sub-universes that are random as well and getting this kind of stability is a great result that illustrates how a top-down approach can extract relevant information and use it to generate stable performance as a strategy.

Conclusion

Handling the market as a network of agents is challenging from many different perspectives. In this paper, we focus on the optimisation of a portfolio that contains a very large number of agents. Traditional optimisation algorithms being unusable, we implement an Enhanced Genetic Algorithm that replicates a Darwinian evolution process. Because of its heavy use of randomness, it is of paramount importance to make sure that the GA is stable enough to be part of an investment strategy.

Our results show the impressive stability of the optimised portfolio when looking at many different financial metrics.

Metric	Mean	Std	Median	1st quartile	3rd quartile
Annualised return	12.54%	0.61%	12.53%	12.16%	12.88%
Annualised volatility	11.85%	0.27%	11.86%	11.70%	12.01%
Risk-reward	1.06	0.05	1.05	1.03	1.08
Percentage up days	52.27%	0.46%	52.28%	51.97%	52.56%
Percentage down days	47.72%	0.45%	47.72%	47.44%	48.03%
Average up days	0.49%	0.00%	0.49%	0.48%	0.49%
Average down days	-0.43%	0.00%	-0.48%	-0.44%	-0.43%
Worst drawdown	-27.86%	1.37%	-28.09%	-28.70%	-27.05%

Table 1: Statistics about the stability of the $SU1$ backtest (100 runs). Past performance and volatility are not a guide to current or future results

Metric	$SU1$	$SU2$	$SU3$
Annualised return [%]	12.39 (0.57)	12.39 (0.71)	12.51 (0.55)
Annualised volatility [%]	11.76 (0.28)	11.84 (0.22)	11.54 (0.40)
Risk-reward	1.05 (0.05)	1.05 (0.06)	1.09 (0.05)
Percentage up days [%]	51.99 (0.47)	52.22 (0.39)	52.37 (0.40)
Percentage down days [%]	48.00 (0.46)	47.77 (0.39)	47.61 (0.41)
Average up days [%]	0.49 (0.00)	0.49 (0.00)	0.47 (0.00)
Average down days [%]	-0.43 (0.00)	-0.43 (0.00)	-0.41 (0.00)
Worst drawdown [%]	-27.85 (1.41)	-27.92 (1.21)	-27.89 (1.30)

Table 2: Comparison between the first 3 sets of random sub universes. Each of them has been run 100 times. Standard deviation of all metrics is in parenthesis. Past performance and volatility are not a guide to current or future results

Metric	$SU4$	$SU5$	$SU6$
Annualised return [%]	12.44 (0.37)	12.76 (0.45)	12.89 (0.55)
Annualised volatility [%]	11.59 (0.22)	11.82 (0.22)	11.67 (0.24)
Risk-reward	1.07 (0.03)	1.08 (0.03)	1.10 (0.05)
Percentage up days [%]	51.94 (0.34)	52.04 (0.28)	52.57 (0.36)
Percentage down days [%]	48.04 (0.34)	47.95 (0.28)	47.42 (0.36)
Average up days [%]	0.48 (0.00)	0.49 (0.00)	0.48 (0.00)
Average down days [%]	-0.43 (0.00)	-0.43 (0.00)	-0.42 (0.00)
Worst drawdown [%]	-27.57 (1.06)	-28.51 (1.42)	-28.53 (1.05)

Table 3: Comparison between the last 3 sets of random sub universes. Each of them has been run 100 times. Standard deviation of all metrics is in parenthesis. Past performance and volatility are not a guide to current or future results

	$SU1$	$SU2$	$SU3$	$SU4$	$SU5$	$SU6$
$SU1$	1.000					
$SU2$	0.999	1.000				
$SU3$	0.990	0.991	1.000			
$SU4$	0.992	0.993	0.998	1.000		
$SU5$	0.991	0.992	0.992	0.994	1.000	
$SU6$	0.990	0.991	0.995	0.995	0.992	1.000

Table 4: Correlation between the average returns of backtests $SU1$ to $SU6$.

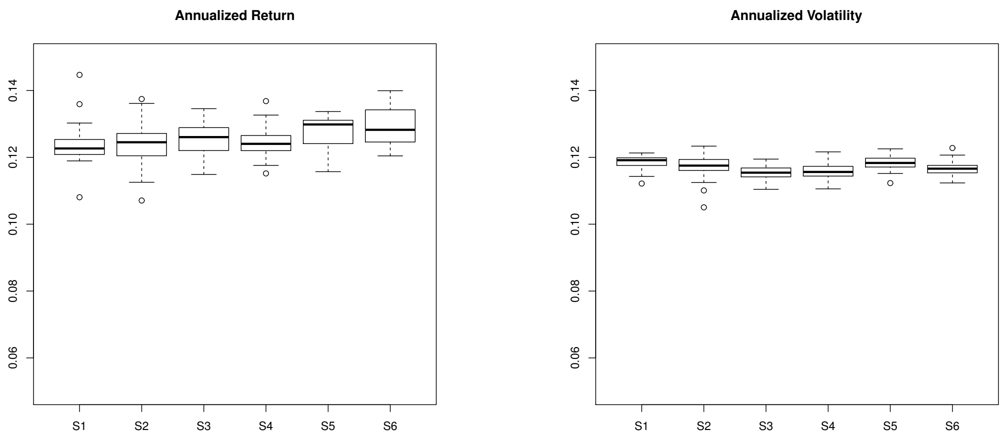

Figure 1: Boxplots of the annualised return and volatility of the 6 datasets. Past performance and volatility are not a guide to current or future results

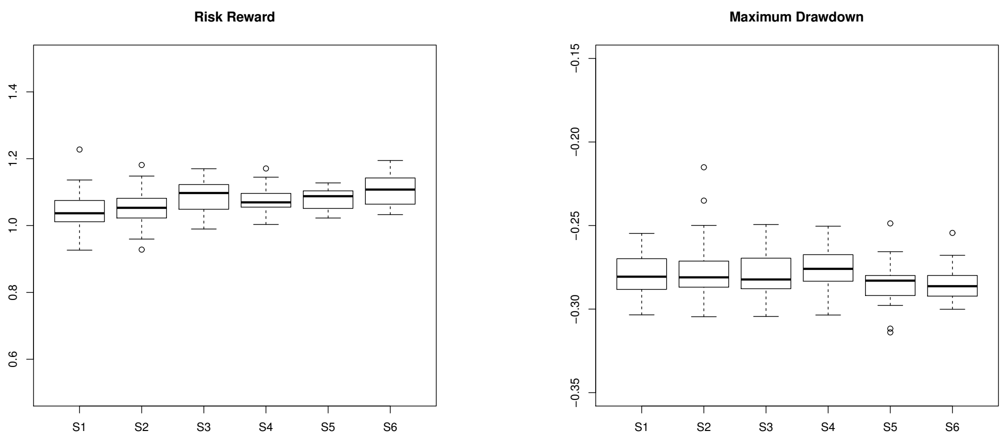

Figure 2: Boxplots of the risk/reward and maximum drawdown of the 6 datasets. Past performance and volatility are not a guide to current or future results

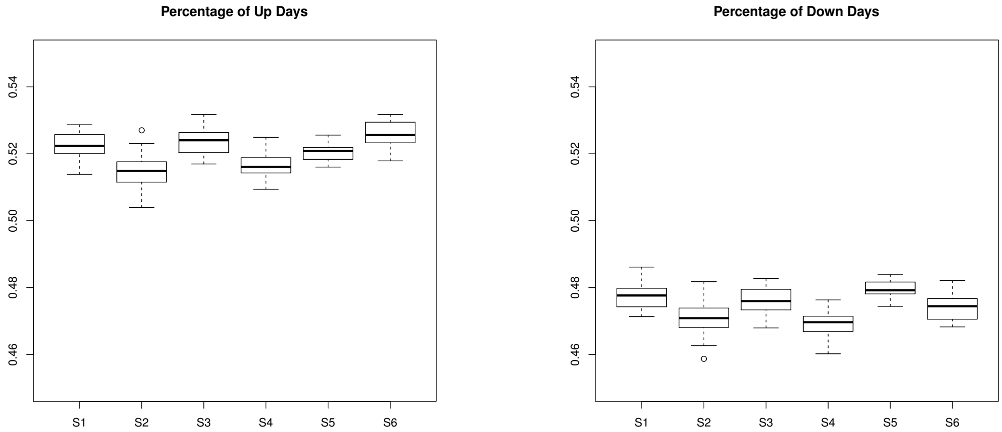

Figure 3: Boxplots of the percentage of up and down days of the 6 datasets. Past performance and volatility are not a guide to current or future results

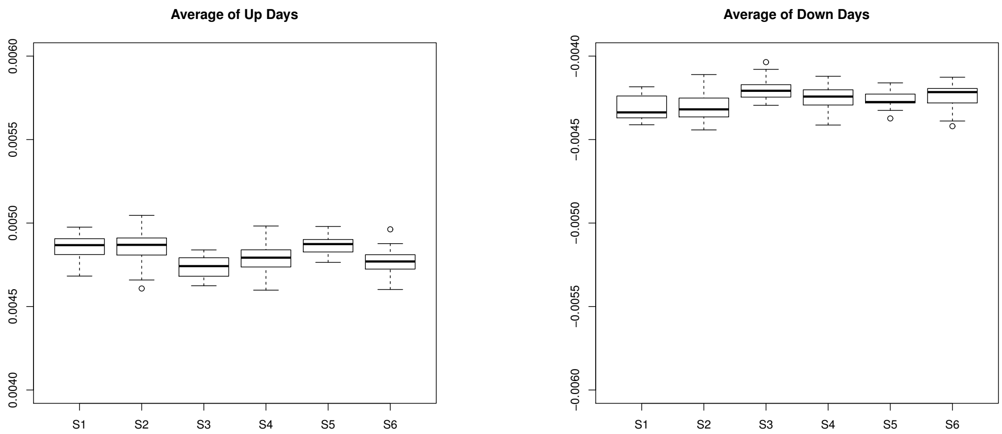

Figure 4: Boxplots of the average daily return of up and down days of the 6 datasets. Past performance and volatility are not a guide to current or future results

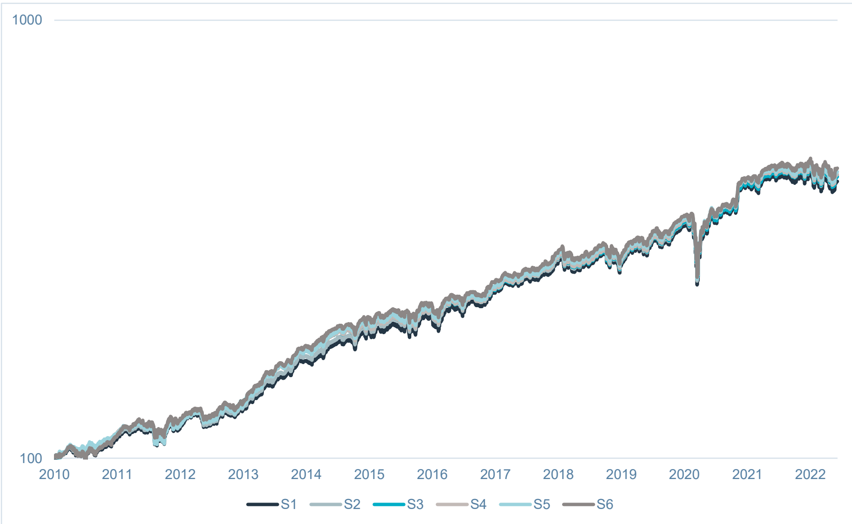

Figure 5: Historical performance of backtests (logarithmic scale). Past performance and volatility are not a guide to current or future results

Disclaimer

This marketing document is only provided for information purposes to professional clients, and it does not constitute an offer, investment advice or a solicitation to subscribe shares in any jurisdiction where such an offer or solicitation would not be authorised or it would be unlawful. Past performance is not a guide to current or future results. There is no guarantee to get back the full amount invested. Particular attention is paid to the contents of this document but no guarantee, warranty or representation, express or implied, is given to the accuracy, correctness or completeness thereof. Issued in Switzerland by RAM Active Investments S.A. which is authorised and regulated in Switzerland by the Swiss Financial Market Supervisory Authority (FINMA). Issued in the European Union and the EEA by the authorised and regulated Management Company, RAM Active Investments (Europe) S.A., 51 av. John F. Kennedy L-1855 Luxembourg, Grand Duchy of Luxembourg.