GENETIC ALGORITHMS: A HEURISTIC APPROACH TO MULTI-DIMENSIONAL PROBLEMS

Written by Philippe Huber and Tony Guida

August 2019

Evolutionary algorithms are not new and have been developed, both their concepts and framework, since around the 1950's based on the idea that the evolutionary process could be used as a general-purpose optimization tool. The goal of this paper is to propose an alternative to classical optimization techniques that can handle systems of a very high dimension. With the rapid rise of computing power, as well as the augmentation of alternative sources of data, quantitative analysts are confronted by numerical challenges that didn't exist a decade ago. In this paper, we show that a Genetic Algorithm (GAs) is a simple process based on the evolution paradigm that is well adapted to very large portfolios, increasing the execution speed; an optimization of a portfolio of more than 100'000 times series of 5'000 daily returns takes less than 5 minutes. Finally, we illustrate that, although GAs are a random process that generates a different solution every time it is run on the same data, it is remarkably stable.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

Introduction

Evolutionary Algorithms are not new and have been developed, both their concepts and framework, since around the 1950's based on the idea that the evolutionary process; a slow change in population through mutation and cross-breeding for the survival of the fittest, could be used as a general-purpose optimization tool. This field of research has been one of many initiatives and desires to model nature and the process underlying phenomena that we have taken for granted. Conceptually based on Darwin's evolutionary theory (Robert Darwin (1859)), Evolutionary Algorithms and especially Genetic Algorithms are based on an attempt to model the process behind natural selection. Evolutionary Algorithms have been considered as a general-purpose heuristic optimization for larger complex systems in different areas of research, from pattern recognition and portfolio selection to bio-engineering.

History

Seminal works to define both a mathematical framework and the theoretical foundations behind the evolutionary process of species were produced around the 1960s. Holland (1994) shows that Genetic Algorithms enabled the exploration of a greater range of potential solutions to a given problem, emulating de facto nature, by creating artificial population and making it evolve through crossbreeding and natural selection. Holland's purpose was to identify and develop a framework where he can formalize the Genetic Algorithm into an evolutionary programming approach for large computational problems. Bremermann's work (J. Bremermann et al. (1966)) focused on added characteristic cross-over elements through the reproduction of parents that were both giving genes to offspring. The "mating" process was not without its flaws since it could be applied only to characteristics that could be added together when following the same population distribution. The interested reader should study Bäck and Schwefel (1993) for a detailed but concise review of the evolutionary strategies and techniques in the field of Evolutionary Algorithms. At this point in time, some of the main operators inspired by evolution were defined as mutation, crossbreeding and natural selection.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

Dimensionality reduction

Due to the heuristic-based nature and its prior objective, Genetic Algorithms appear as one extremely well-suited and innovative optimization method to deal with a high dimensionality case where traditional optimisation methods often fail, as evidenced in Raymer et al. (1997). In this paper, they apply GAs in feature selection and extraction as a means to reduce the high dimensionality and to improve the accuracy of a classifier in a pattern recognition problem using a medical dataset on hypothyroid population. The use of a GA allows to simultaneously perform features selection and features transformation, benefiting from the cross-over and mutation operators. This feature selection approach was first introduced by Siedlecki and Sklansky (1989). The effective use of parallelism ensures the possibility to simultaneously explore different combinations and find an efficient solution in a vast set of possible opportunities, which has been perfectly suited for dimensionality reduction purposes. Comparisons of commonly accepted dimensionality reduction techniques have been researched by Nick et al. (2015). In their papers, they compare principal component analysis (PCA), recursive feature elimination and genetic and evolutionary weighting and selection (GEFeWS) before training an SVM classifier. Results show that the genetic-based technique produced better accuracy than traditional PCA. Naturally, this appealing set of interesting characteristics promoted the use of Genetic Algorithms into a large number of computational use-cases coming from extremely different industries: from healthcare to aircraft modeling or bio-engineering in the search of a protein design for drug creation.

Genetic Algorithms and finance

In the field of computational algorithms, the term "search space" is used to qualify the many potential solutions for a specific problem. Naturally, an optimization tool that can efficiently, adaptively and rapidly scan through millions of potential portfolios, became highly appealing for finance professionals. In 2010, Torrubiano and Suarez (2010) used a hybrid method combined with a Genetic Algorithm for selecting constrained portfolios, that traditional mixed-integer quadratic optimization would solve with some difficulty. Genetic Algorithms, benefiting from their heuristic nature, could solve portfolio construction problems that traditional methods could not, by reducing the dimension of the optimization problem and removing portfolios that are not "fit" to survive. Using Genetic Algorithms is not a guarantee of finding a near optimum in the optimization process, but the proba-

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

bility is higher relative to using a traditional quadratic. Evolutionary Algorithms require no fitness gradient information of any kind to proceed, are easy to process in parallel and have the ability to escape from local minima where deterministic optimization methods may fail or are not applicable for non-deterministic polynomial acceptable problems.

Genetic Algorithms and AI

Literally coding rules by hand and specifying in advance the parameters of the models are more linked to the usual modeling approach still taught in University. The Galilean approach states that a model should pass through the flames of experiments and observations for the results. Soon the issue of performing a specific task in a non-variable environment became clear, hence the need for a computational paradigm that allows for more adaptive modeling in a dynamic environment. How to confer intelligence to a program and make sure that the program will learn without being taught to, or without having the rules formally written and explained. To those regards one can see where Artificial Intelligence models could be related to Genetic Algorithms by certain axioms as: adaptive computer program, optimization for high dimensional problems, computationally efficient using parallelism and obviously their common evolutionary aspect when we focus on artificial neural nets. Artificial Neural Networks (ANNs) like GAs are biologically motivated approaches. Indeed, most machine learning (ML) models could summarize by an optimization under some regularization. However, this is where the family links stops. Genetic Algorithms remain very different, GA's utilize a heuristic approach, whereas ML minimises a loss using a partial derivative on the gradients. The modeling using features according to fit/label in a supervised manner via a statistical learning model is different by nature compared to the GA approach. The current link and plausible future for the interactions between Genetic Algorithms and artificial neural nets would be in the feature selection and transformation (E. Goldberg (1989)), and in the parameter's optimization, more lately called hyperparameters tuning or automatic machine learning, which usually are based using Bayes priors (Feurer et al. (2015)). GAs could be used in different steps of a neural net algorithm, through features selection (Raymer et al. (2000)), to replacing back propagation (J. Montana and Davis (1989)) and potentially evolutionary reinforcement learning.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

Data

Our proprietary dataset consists of 74 characteristics of 912 equities listed on US markets. The list of features are provided in Table 4. The vast majority of features are accounting-based and proxy a large scope of documented anomalies: size (F. Fama Sr and French (1992), van Dijk (2011), Asness et al. (2018)), value (F. Fama Sr and French (1992), Pätäri and Leivo (2015)), profitability and investment (F. Fama and French (2014))) and quality (S. Asness et al. (2014)). Some features are price-based (past returns) and reflect momentum-like patterns (Novy-Marx (2012), S. Asness et al. (2014)) or volatility-based (related to low-risk strategies as in Baker et al. (2013) and the references therein). Most equities being US-based, features are quoted in USD. The chronological range is January 2000 to December 2018, and the points are sampled on a monthly frequency. We engineer all features so that, each month, their value is equal to their level on the empirical cumulative distribution function of the current month. The features are thus quantile scores and, as such, all predictor values lie inside the unit interval and are uniformly distributed for a given month. We recall that normalizations are commonplace, both in portfolio selection (e.g., W. Brandt et al. (2009) or Ammann et al. (2016)) and in the asset pricing literature (T. Kelly et al. (2019), S.J. Koijen and Yogo (2018)). Given the amount of data at our disposal, providing simple descriptive statistics is impractical.

Why a Genetic Algorithm?

With these data, we build simple strategies by applying several mathematical transformations (like moving averages for instance) on all our variables. Those transformations can have 1, 2 or 3 periods as parameters. On each trading day, all stocks are ranked and strategies will take short and long positions based on these rankings.

Here, we make no *a priori* assumptions: we show no preference between short term or long term strategies, or between mean reverting and trend following strategies. Therefore, we run a grid search on the transformations parameters over a large range of values going from 2 days to 1000 days. All strategies with a negative average return over the in-sample period are eliminated. This leads to a set of about $N = 100,000$ strategies which are represented by time series of length T. Optimizing a portfolio of N assets under the Markowitz paradigm is performed through the solving of the quadratic program

$$\begin{aligned} \underset{\mathbf{w}}{\text{maximize}} \quad & \alpha^T \mathbf{w} - \lambda \mathbf{w}^T Q \mathbf{w} \\ \text{subject to} \quad & E\mathbf{w} = \mathbf{d} \end{aligned}$$

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

where α is the vector of expected returns, Q is the covariance matrix of returns and $\lambda > 0$ is a parameter describing the investor's risk aversion. When the problem is properly conditioned, i.e. when the matrix Q is positive definite, classical algorithms like conjugate gradient or interior point are efficient and easy to use, as long as N stays small. In the present study, we are dealing with a dataset of $n \approx 100,000$ time series of daily data with a length of approximatively $T = 5,000$ days. This brings two challenges that make the use of a quadratic program impossible:

1. The rank of the covariance matrix is equal to $T \lll N$, hence Q is not positive definite,

2. N is much too big.

Therefore, we propose to use a Genetic Algorithm (GA) that, thanks to its metaheuristic approach, is able to generate high-quality solutions in a tractable computing time for a problem of this size. Indeed, a GA does not require the computation of partial derivatives or a covariance matrix which have a computational complexity $\mathcal{O}(TN^2)$; only the objective function is needed ($\mathcal{O}(TN)$).

How does a Genetic Algorithm work?

Genetic Algorithms belong to the group of Evolutionary Algorithms and are inspired by Darwin's theory of evolution. They follow a metaheuristic process which means that they are non-deterministic. In other words, a GA uses a space search to find a near-optimal solution by running a random procedure.

There are several ways of setting up a GA and, in this paper, we decided to follow these 5 steps:

1. **Initialization**: a first population of P candidate solutions (chromosomes) are randomly generated (Generation 1). Each chromosome contains N parameters (genes),

2. **Evaluation**: each chromosome is evaluated by an objective function f,

3. **Selection**: among all chromosomes, two are selected because they have the best objective function values,

4. **Crossover**: they become the parents of the next generation of N chromosomes. Each gene of the offspring is the exact replication of either the mother's or the father's

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

gene. In addition, a random mutation is added to the gene with low probability to avoid getting stuck in a local optimum,

5. **Repeat** from step 2 until some criterion is met.

Application to a high dimensional portfolio

In this section, we illustrate our use of Genetic Algorithms to build a portfolio with a large set of systematic strategies.

As stated in the previous two sections, optimizing a portfolio of this size is not possible with a classical algorithm and for this reason, we have utilized a go for the Genetic Algorithm that we described hitherto.

The objective function is the total return of the portfolio and is maximized with regards to a vector of weights $\mathbf{w}$ of length N. Weights are constrained to be non negative (strategies cannot be shorted) and the volatility of the portfolio σ must be smaller than an upper bound σ_{max}:

$$\begin{aligned} \underset{\mathbf{w}}{\text{maximize}} \quad & f(\mathbf{w}) = \prod_{t=1}^{T} \left(1 + \sum_{i=1}^{N} w_i r_{it}\right) - 1 \\ \text{subject to} \quad & w_i \geq 0,\ i = 1, \ldots, N \\ & \sigma \leq \sigma_{max}. \end{aligned}$$

where r_{it} is the return of strategy i on day t.

In the first step of the algorithm, a population of P portfolios $p_1 \ldots p_P$ is generated with random weights $\mathbf{w_1} \ldots \mathbf{w_P}$. Each portfolio is evaluated with the objective function f and then ranked. The best two portfolios become parents for Generation 2 (as described in step 4 of the previous section). The process is repeated until a stopping criterion is met; here, we define a maximum number of generations I_{max}. The whole process is illustrated by the Algorithm 1.

In order to produce a meaningful simulation, the process is run on a yearly rolling basis so that it can adapt over time: we use data from 2000 to 2008 to simulate a portfolio in 2009, then data from 2000 to 2009 to simulate 2010, and so on until 2018. Costs and slippage are included in the results. The number of strategies for each year of simulation is shown in Table 1.

Regarding computing times, Genetic Algorithms offer the advantage of being fully parallelizable and, therefore, are particurlarly well suited to be coded on GPU cards. Convergence is very fast, considering the high dimensionality of the problem (Table 2): on

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

Algorithm 1 Genetic Algorithm

```
Generation 1:
    for i := 1, i ≤ P do
        for j := 1, j ≤ N do
            w_ij = random value
        end for
        evaluate each portfolio P_i by computing f(w_i)
    end for
    sort(P_1, ..., P_P) w.r.t f and store in (P̃_1, ..., P̃_p)
    store P̃_1 and P̃_2 as parents of Generation 2

Generations 2 to I_max:
    w̃_1j := weight j of Parent P̃_1
    w̃_2j := weight j of Parent P̃_2
    for i := 1, i ≤ P do
        for j := 1, j ≤ N do
            uniform crossover:
                generate r := random_uniform(0, 1)
                if r > 0.5 then
                    w_ij := w̃_1j
                else
                    w_ij := w̃_2j
                end if
            mutation:
                generate r := random_uniform(0, 1)
                if r < 0.001 then
                    generate s := random_normal(0, 0.0001)
                    w_ij := w_ij + s
                end if
        end for
        evaluate each portfolio P_i by computing f(w_i)
    end for
    sort(P_1, ..., P_P) w.r.t f and store in (P̃_1, ..., P̃_p)
    Store P̃_1 and P̃_2 as parents of the next Generation

return P̃_1 of Generation I_max as the optimal portfolio
```


The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

In-sample period	Year	Strategies
2000 - 2008	2009	102,920
2000 - 2009	2010	92,380
2000 - 2010	2011	93,620
2000 - 2011	2012	62,465
2000 - 2012	2013	68,758
2000 - 2013	2014	83,390
2000 - 2014	2015	91,512
2000 - 2015	2016	97,836
2000 - 2016	2017	123,256
2000 - 2017	2018	137,454

Table 1: Number of strategies in the portfolio for each year of simulation.

Year	Strategies	Time
2009	102,920	163s
2010	92,380	145s
2011	93,620	154s
2012	62,465	108s
2013	68,758	114s
2014	83,390	137s
2015	91,512	142s
2016	97,836	161s
2017	123,256	208s
2018	137,454	231s

Table 2: Computing time on a single P100 Nvidia GPU.

average, 3 minutes are enough to optimize systems of around 100,000 time series with a single GPU card.

Simulation results are displayed on Figure 1 and Table 3. The goal of this paper is not to propose a set of strategy outperforming the market. Nevertheless, we see that, even with very basic signals and factors, the possibility of handling a large number of different strategies provides interesting results.

More importantly, we show that simulation results are stable when the whole process is run several times. We performed 100 simulations and display averages, volatilities as well as information ratios of returns with their standard deviations in parenthesis in Table 3. The tightness (low standard deviation) of each boxplot is remarkable and confirms that solutions are stable. Figure 3 illustrates it with the distribution of the time series of returns

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

Year	Return	Volatility	IR
2009	-3.58% (0.93%)	5.47% (0.19%)	-0.66 (0.17)
2010	-3.16% (0.71%)	4.15% (0.08%)	-0.76 (0.17)
2011	-5.35% (0.85%)	4.20% (0.09%)	-1.27 (0.21)
2012	-1.14% (0.60%)	3.53% (0.07%)	-0.32 (0.17)
2013	6.40% (0.61%)	3.80% (0.09%)	1.69 (0.15)
2014	3.98% (0.85%)	4.58% (0.09%)	0.87 (0.18)
2015	6.90% (0.72%)	5.92% (0.13%)	1.17 (0.11)
2016	6.07% (0.75%)	5.12% (0.11%)	1.19 (0.15)
2017	0.93% (0.96%)	4.17% (0.11%)	0.22 (0.23)
2018	0.92% (0.88%)	3.62% (0.10%)	0.25 (0.24)
All*	1.20% (0.26%)	4.52% (0.04%)	0.26 (0.06)

Table 3: Simulated returns, volatilities and information ratios with standard deviations in parenthesis. *Numbers for the whole period are annualized.

for all 100 simulations.

Conclusion

The goal of this paper was to propose an alternative to classical optimization techniques that can handle systems of very high dimensions. With the rapid rise of computing power as well as the augmentation of alternative sources of data, quantitative analysts are confronted by numerical challenges that didn't exist a decade ago. We showed that a Genetic Algorithm is a simple process based on the evolution paradigm that is well adapted to very large portfolios. In addition, it is very fast since an optimization of a portfolio of more than 100'000 times series of 5'000 daily returns takes less than 5 minutes. Finally, we illustrated that although GAs are a random process that generates a different solution every time it is run on the same data, it is remarkably stable. With the augmentation of available data, heuristic optimization algorithms will likely become a standard in the near future. Genetic Algorithms in particular are powerful and well suited for parallel programming.

Signal	Description
adv_3m	average daily volume over the last 3 months
adv_6m	average daily volume over the last 6 months
asset_turnover	company revenue on assets
bps_1yg	book per share 1 year growth
bv	book value
capex_ps_cf	capital expenditures per share

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

cash_div_cf	cash dividends/cash flow
cash_per_share	cash per share
com_eq_assets	common equity % total assets
com_eq_tcap	equity % total capital
cps_1yg	cash per share one year growth
debt_assets	total debt % total assets
debt_eq	total debt % total equity
dep_accum_fix_assets	accumulated depreciation % gross fixed assets
div	dividend paid total
div_yield	dividend yield
dps_1yg	dividend per share one year growth
ebit_ta	ebit on total asset
eps	earnings per share - fiscal period
eps_basic	eps - basic - before extraordinaries
eps_dil	eps - fully diluted
eps_dil_bef_unusual	eps - diluted - before unusual expense
eps_dil_gr	eps - diluted - before extras - % change
eps_secs	earnings per share - security
eps_xord	earnings per share incl. extraordinary items - fiscal
eq_1yg	equity one year growth
fcf	free cash flow
fcf_oa	free cash flow on operating asset
fcf_ta	free cash flow on total asset
ffo	fund from operations
fix_assets_com_eq	fixed assets % common equity
free_ps_cf	cash flow per share (diluted) - free
gp_oa	gross profit on operating asset
gp_ta	gross profit on total asset
ins_rev_1yg	insurance reserves 1 year growth
intang	intangibles
min_int_tcap	minority interest % total capital
mkt_cap_12m	market cap 12 months average
mkt_cap_3m	market cap 3 months average
mkt_cap_6m	market cap 6 months average
mom_11m_lcl	price momentum 12-1 months
mom_5m_lcl	price momentum 6-1 months
mom_sharp_11m_lcl	price mom 12-1 divided by price volatility
mv_1yg	market value growth
net_cf_debt	cash flow (net) % total debt
net_margin	net margin
ni_1yg	net income one year growth
oa	operating asset
ocf	operating cash flow

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

ocf_bv	operating cash flow on book value
ocf_margin	operating cash flow margin
ocf_oa	operating cash flow operating asset
oi	operating income
oi_margin	operating margin
oper_mgn	operating margin
ptx_mgn	pretax margin
r12m_lcl	12 months return local currency
r18m_lcl	18 months return local currency
r1d_lcl	1 day return local currency
r1m_lcl	1 month return local currency
r36m_lcl	36 months return local currency
r3m_lcl	3 months return local currency
r6m_lcl	6 months return local currency
r9m_lcl	9 months return local currency
recurring_ear_tot_asset	recurring earning on total assets
roa	return on asset
roc	return on capital
s_1yg	sales 1 year growth
sales_ps	sales per share
std_debt	short term debt % total debt
total_liab_total_asset	total liabilities on total asset
vol1y_lcl	price volatility one year
vol3y_lcl	price volatility 3 years
vol5y_lcl	price volatility 5 years

Table 4: List of features used in simulations.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

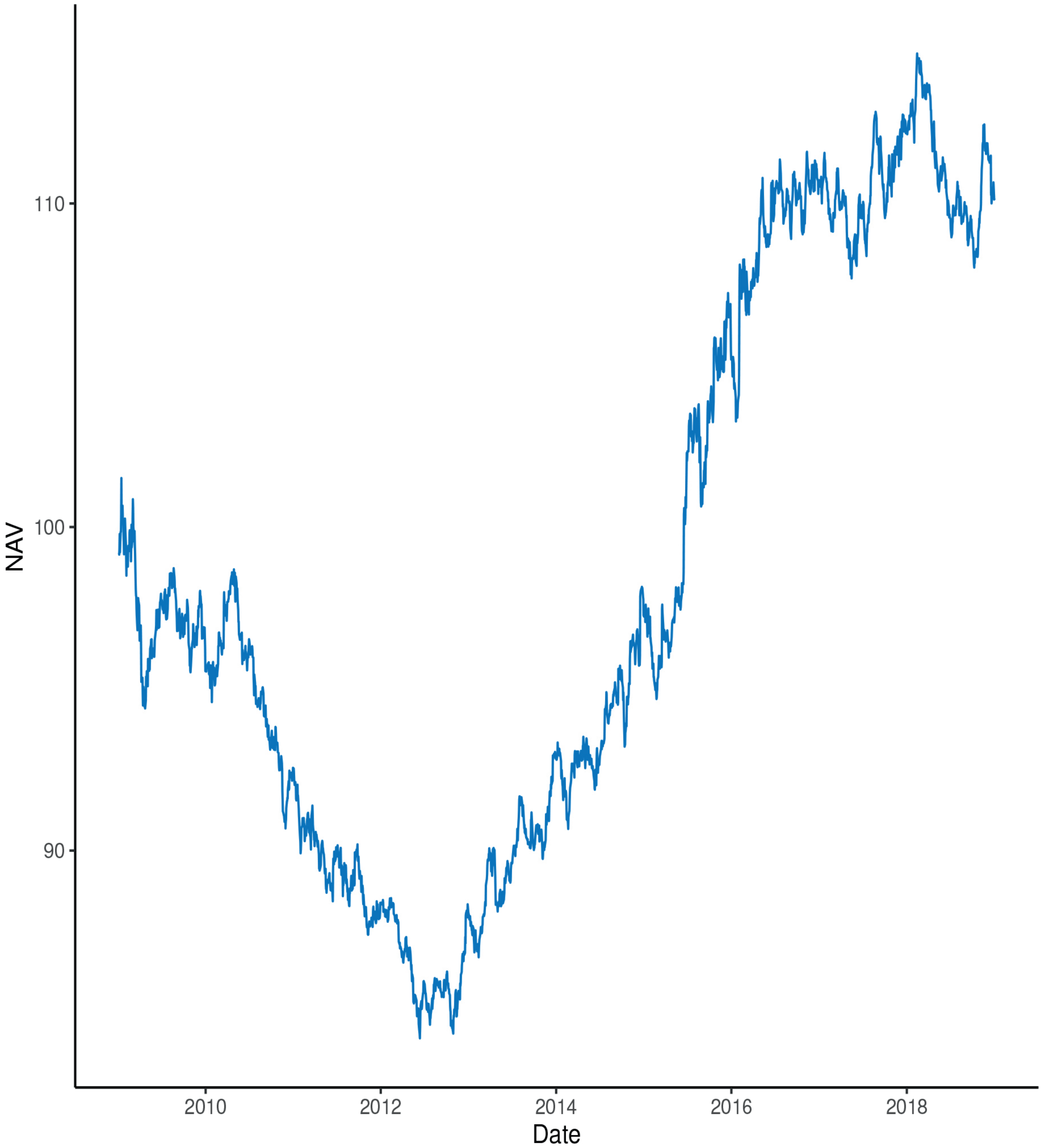

Figure 1: Chart of a simulated portfolio of strategies optimized by a Genetic Algorithm

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

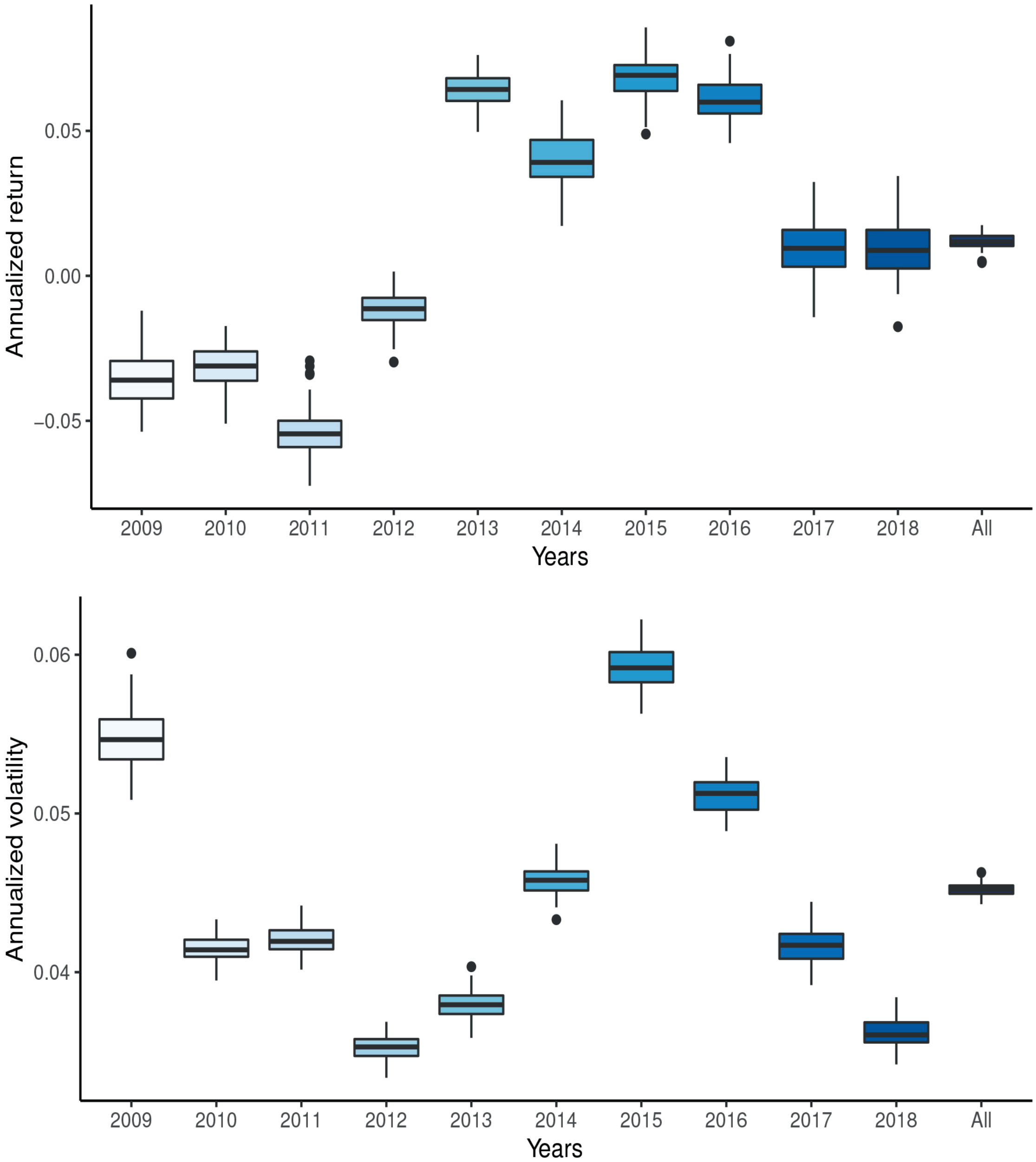

Figure 2: Boxplots of the returns and volatility of the optimal portfolio by year and for the whole period

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

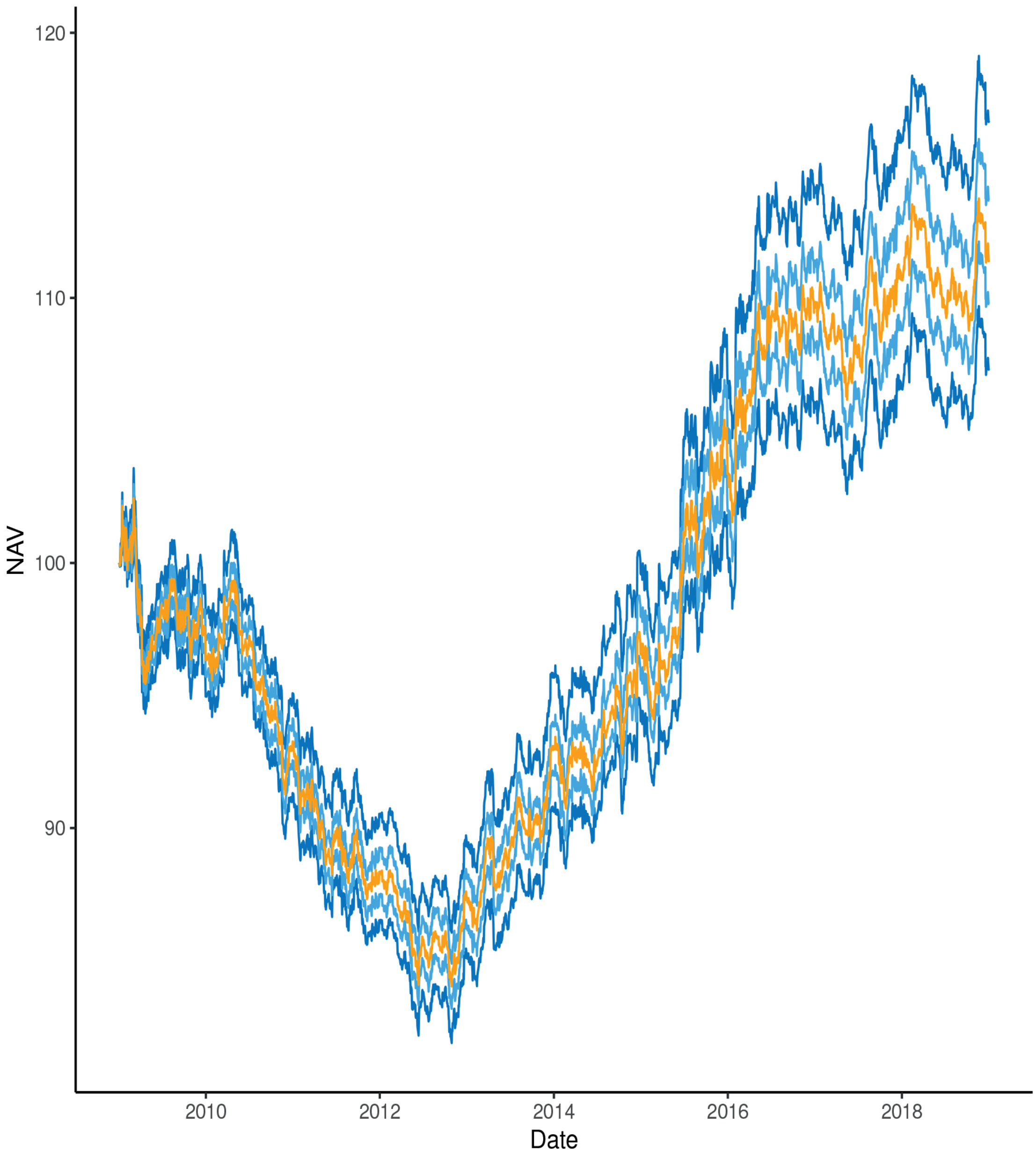

Figure 3: Distribution of the optimized portfolio returns. Dark blue = 5% and 95% percentiles, light blue = 25% and 75% quartiles, orange = median.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

References

Ammann, M., Coqueret, G., and Schade, J.-P. (2016). Characteristics-based portfolio choice with leverage constraints. *Journal of Banking And Finance*, 70.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T., and Pedersen, L. (2018). Size matters, if you control your junk. *Journal of Financial Economics*, 129.

Bäck, T. and Schwefel, H.-P. (1993). An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1:1–23.

Baker, M., Bradley, B., and Taliaferro, R. (2013). The low risk anomaly: A decomposition into micro and macro effects. *Financial Analysts Journal*, 70.

E. Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization*, volume 2.

F. Fama, E. and French, K. (2014). A five-factor asset pricing model. *Journal of Financial Economics*, 116.

F. Fama Sr, E. and French, K. (1992). The cross-section of expected stock return. *Journal of Finance*, 47:427–65.

Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., and Hutter, F. (2015). Efficient and robust automated machine learning. In Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems 28*, pages 2962–2970. Curran Associates, Inc.

Holland, J. (1994). Adaptation in natural and artificial systems. *An Introductory Analysis with Application to Biology, Control and Artificial Intelligence*.

J. Bremermann, H., Rogson, M., and Salafi, S. (1966). Global properties of evolution processes. *Evolutionary Computation: The Fossil Record*.

J. Montana, D. and Davis, L. (1989). Training feedforward neural networks using genetic algorithms. pages 762–767.

Nick, W., Shelton, J., Bullock, G., Esterline, A., and Asamene, K. (2015). Comparing dimensionality reduction techniques. *Conference Proceedings - IEEE SOUTHEASTCON*, 2015.

Novy-Marx, R. (2012). Is momentum really momentum. *Journal of Financial Economics - J FINAN ECON*, 103.

Pätäri, E. and Leivo, T. (2015). A closer look at value premium: Literature review and synthesis. *Journal of Economic Surveys*, 31:n/a–n/a.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

Raymer, M., Punch, W., Goodman, E., Kuhn, L., and K. Jain, A. (2000). Dimensionality reduction using genetic algorithms. *Evolutionary Computation, IEEE Transactions on*, 4:164 – 171.

Raymer, M., Punch, W., Goodman, E., Sanschagrin, P., and Kuhn, L. (1997). Simultaneous feature scaling and selection using a genetic algorithm. pages 561–567.

Robert Darwin, C. (1859). On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. *London: J. Murray*, 23.

S. Asness, C., Frazzini, A., Israel, R., and Moskowitz, T. (2014). Fact, fiction and momentum investing. *SSRN Electronic Journal*, 40.

Siedlecki, W. and Sklansky, J. (1989). A note on genetic algorithms for large-scale feature selection. *Pattern Recognition Letters*, 10:335–347.

S.J. Koijen, R. and Yogo, M. (2018). A demand system approach to asset pricing. *Journal of Political Economy*.

T. Kelly, B., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*.

Torrubiano, R. and Suarez, A. (2010). Suarez, a.: Hybrid approaches and dimensionality reduction for portfolio selection with cardinality constraints. ieee computational intelligence magazine 5(2), 92-107. *Computational Intelligence Magazine, IEEE*, 5:92 – 107.

van Dijk, M. (2011). Is size dead? a review of the size effect in equity returns. *Journal of Banking And Finance*, 35.

W. Brandt, M., Santa-Clara, P., and Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, 22:3411–3447.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

The views expressed in this document are the author's and do not necessarily reflect those of the organizations he is affiliated with. The content of this document is intended for information purpose only and is not an official statement of RAM AI SA. No investment decision or particular course of action is recommended by this presentation from the author.

GENETIC ALGORITHM: AN HEURISTIC APPROACH TO MULTI-DIMENSIONAL PROBLEMS

Written by Philippe Huber and Tony Guida

August 2019