Dead Alphas as Risk Factors

Zura Kakushadze $\S^{\dagger 1}$ and Willie Yu $^{\sharp 2}$

§ Quantigic® Solutions LLC
1127 High Ridge Road #135, Stamford, CT 06905 ³

† Free University of Tbilisi, Business School & School of Physics
240, David Agmashenebeli Alley, Tbilisi, 0159, Georgia

‡ Centre for Computational Biology, Duke-NUS Medical School
8 College Road, Singapore 169857

(July 31, 2017)

Abstract

We give an explicit algorithm and source code for extracting equity risk factors from dead (a.k.a. "flatlined" or "hockey-stick") alphas and using them to improve performance characteristics of good (tradable) alphas. In a nutshell, we use dead alphas to extract directions in the space of stock returns along which there is no money to be made (and/or those bets are too volatile). In practice the number of dead alphas can be large compared with the number of underlying stocks and care is required in identifying the aforesaid directions.

¹ Zura Kakushadze, Ph.D., is the President of Quantigic[®] Solutions LLC, and a Full Professor at Free University of Tbilisi. Email: zura@quantigic.com

 $^{^2\,}$ Willie Yu, Ph.D., is a Research Fellow at Duke-NUS Medical School. Email: willie.yu@duke-nus.edu.sg

³ DISCLAIMER: This address is used by the corresponding author for no purpose other than to indicate his professional affiliation as is customary in publications. In particular, the contents of this paper are not intended as an investment, legal, tax or any other such advice, and in no way represent views of Quantigic[®] Solutions LLC, the website www.quantigic.com or any of their other affiliates.

1 Introduction and Summary

In the olden days quantitative trading (e.g., statistical arbitrage) strategies typically were based on some kind of an underlying idea that had to "make sense". Because of this, the number of trading strategies back then (excluding evident variations) was relatively limited. But then hardware got really cheap and quant traders turned to data mining⁴ which resulted in exponential proliferation of alphas⁵ in modern quant trading,⁶ and these ever-ubiquitous alphas are ever fainter and more ephemeral.

Alphas die – that is, "flatline" or become so faint that they are no longer tradable (a.k.a. "hockey-stick" alphas) – and are replaced with new data-mined alphas all the time, some or all of which in time will likely suffer the same fate. Typically, one simply tosses these dead alphas. In this note we ask the following question:

Can we use the information encoded in dead alphas to improve the performance of good (tradable) alphas? Put differently, can we learn something from dead alphas that we could possibly use to make good alphas more profitable and/or less volatile, etc.? As we discuss in the following, the answer to this question is affirmative. Thus, dead alphas define directions in the space of stock returns along which there is no money to be made (and/or those bets are too volatile). And this is valuable information. We can use it to make our good alphas even better. How? The idea here is simple. We can factor out these directions from good alphas, i.e., tweak them such that they do not make bets in these unproductive directions even partially.

In terms of a practical implementation, to identify the aforesaid directions in the space of stock returns associated with dead alphas we use the underlying stock position data, which is known for both dead and good alphas. What we end up extracting from this data is essentially a risk factor loadings matrix using which we can neutralize (via, e.g., a regression) stock positions of good alphas or build a multifactor risk model for optimizing our stock portfolio. This stock portfolio is obtained i) either via combining good alphas with some weights (see, e.g., [Kakushadze and Yu, 2017a]), or ii) by calculating the expected returns for stocks using the expected returns for good alphas as in [Kakushadze and Yu, 2017c] and directly trading a portfolio of stocks based on these expected returns (without combining alphas).

In Section 2 we give an explicit algorithm for computing the directions associated with dead alphas and factoring them out of good (tradable) alphas. We give the source code for this algorithm in Appendix A.⁸ We briefly conclude in Section 3.

⁴ See, e.g, [Tulchinsky et al, 2015].

⁵ Here "alpha" – following the common trader lingo – generally means any reasonable "expected return" that one may wish to trade on and is not necessarily the same as the "academic" alpha. In practice, often the detailed information about how alphas are constructed may not even be available, e.g., the only data available could be the position data, so "alpha" then is a set of instructions to achieve certain stock (or some other instrument) holdings by some times t_1, t_2, \ldots

⁶ See, e.g., [Kakushadze and Tulchinsky, 2016], [Kakushadze, 2016].

⁷ Our approach is agnostic to the nature of the underlying tradables, which need not be stocks.

⁸ The source code given in Appendix A hereof is not written to be "fancy" or optimized for speed or in any other way. Its sole purpose is to illustrate the algorithms described in the main

2 (Dead) Alphas

Our discussion below a priori is oblivious to the underlying instruments. In practice, the most important application is in the case of equity portfolios comprised of a large number of stocks (e.g., 2,000+ most liquid U.S. stocks). So, for definiteness let us focus on alphas that trade (largely) overlapping portfolios of US stocks. Let the number of all stocks traded be M, and let the number of alphas be $N \gg M$. Let the realized returns for alphas be ρ_{is} (i = 1, ..., N), and the realized returns for stocks be R_{As} (A = 1, ..., M), where the index s = 1, ..., T labels trading days (for definiteness, let s = 1 correspond to the most recent date). Then we have

$$\rho_{is} = \sum_{A=1}^{M} P_{iAs} R_{As} \tag{1}$$

Here on each day labeled by s for each alpha labeled by i the quantities P_{iAs} are nothing but the properly normalized stock positions in the corresponding alpha portfolio. The normalization condition is given by

$$\sum_{A=1}^{M} |P_{iAs}| = 1 \tag{2}$$

There may be additional *linear constraints* on stock positions stemming from all alphas being subject to the same risk management restrictions. E.g., suppose all alphas are dollar-neutral. Then we have:

$$\sum_{A=1}^{M} P_{iAs} \equiv 0 \tag{3}$$

More generally, we can have p linear constraints (typically, $p \ll M$, so we assume that this holds)

$$\sum_{A=1}^{M} P_{iAs} \ Q_{A\alpha} \equiv 0, \quad \alpha = 1, \dots, p$$
 (4)

Without loss of generality we can assume that: i) the p columns of the matrix $Q_{A\alpha}$ are linearly independent;⁹ and ii) the constraints (4) do not imply that $P_{iAs} \equiv 0$ for any given value of A (or else none of the alphas trade the stock labeled by A and we can simply drop it out of the universe). Furthermore, to keep it simple, we will assume that if there is a subset of N_1 alphas $(N_1 < N)$ for which – but not for all alphas – we have additional linear constraints other than (4), then $N_1 \ll N$.

Not all alphas are viable for trading. Thus, suppose we take some time horizon, e.g., a period comprised of d trading days. Then we can define expected returns for

text in a simple-to-understand fashion. Some important legalese is relegated to Appendix B.

⁹ The constraints (4) are invariant under SO(p) rotations $Q \to Q U$, where $U_{\alpha\beta}$ is an orthogonal matrix: $UU^T = U^T U = 1$.

alphas, call them η_{is} , via moving averages based on prior d days' worth of realized returns:¹⁰

$$\eta_{is} = \frac{1}{d} \sum_{s'=s+1}^{s+d} \rho_{is'} \tag{5}$$

Ideally, we want these expected returns to be i) positive, ii) not too low, and iii) not too volatile. We can quantify ii) and iii) as follows. E.g., we can demand that $\eta_{is} \geq \eta_{min}$, where η_{min} is some lower bound on the expected returns below which alphas are no longer worth trading. Also, e.g., we can demand that the Sharpe ratios $\eta_{is}/\sigma_{is} \geq S_{min}$, where S_{min} is some lower bound below which alphas are too volatile. And so on. Here $(\text{Var}(\cdot, d))$ is a d-day moving serial variance):

$$\sigma_{is}^2 = \text{Var}(\eta_{is}, d) = \frac{1}{d-1} \sum_{s'=s+1}^{s+d} (\eta_{is'} - \overline{\eta}_{is})^2$$
 (6)

$$\overline{\eta}_{is} = \frac{1}{d} \sum_{s'=s+1}^{s+d} \eta_{is'} \tag{7}$$

So, we can define dead alphas as those for which, e.g., $\eta_{is} < \eta_{dead}$ and $\eta_{is}/\sigma_{is} < S_{dead}$, where η_{dead} and S_{dead} can – but need not – be the same as η_{min} and S_{min} . Then, similarly, good (tradable) alphas are those for which $\eta_{is} \geq \eta_{min}$ and $\eta_{is}/\sigma_{is} \geq S_{min}$.

2.1 Dead Alphas as Risk Factors

Let us denote the subset of dead alphas as J', i.e., for $i \in J'$ the corresponding alpha is dead. We are not going to trade these alphas. However, simply tossing them would be wasteful too. We can use them to improve good alphas we are trading – let the subset of these alphas be J. We then have $N_{dead} + N_{good} \leq N$, where $N_{dead} = |J'|$ is the number of dead alphas, while $N_{good} = |J|$ is the number of good alphas.¹¹

Dead alphas define directions in the space of stock returns along which there is no money to be made (and/or those bets are too volatile). This is valuable information. We can use it to make our good alphas even better. How? The idea here is simple. We can factor out these directions from good alphas. I.e., we can tweak good alphas such that they do not make bets in these unproductive directions even partially.

However, first we must identify these directions. The issue is that, precisely because of proliferation and ephemeral nature of alphas, typically, the number of dead alphas N_{dead} is large: $N_{dead} \gg M$. It is clear that we cannot factor out more directions than we have stocks. So, then, how can we define these directions?

Here is a simple solution. Let us define the $M \times M$ matrix (for each value of s):

$$X_{ABs} = \sum_{i \in J'} P_{iAs} \ P_{iBs} \tag{8}$$

We emphasize that this is only an example and there are other ways of constructing η_{is} .

We are assuming that $\eta_{dead} \leq \eta_{min}$ and $S_{dead} \leq S_{min}$. Also, in principle, some dead alphas may start performing well again down the road, so one should bear this in mind.

In the presence of the linear constraints (4), this matrix is singular; otherwise, if $N_{dead} \gg M$, generically it is nonsingular. (In the remainder of this paragraph, to simplify notations, we suppress the index s.) We can compute the first K principal components $V_A^{(a)}$, $a=1,\ldots,K$, of X_{AB} (with the corresponding eigenvalues in the descending order, $\lambda^{(1)} > \lambda^{(2)} > \cdots > \lambda^{(K)}$), where we have to determine what to take for K. A simple method is to take K to be the effective rank (or eRank)¹² of the matrix X_{AB} [Roy and Vetterli, 2007]. More precisely, since eRank is fractional, we can take, e.g., $K = \text{Round}(\text{eRank}(X_{AB}))$ or $K = \text{floor}(\text{eRank}(X_{AB})) = \lfloor \text{eRank}(X_{AB}) \rfloor$.

So, restoring the index s, we have K_s principal components $V_{As}^{(a)}$, $a=1,\ldots,K_s$, which define the directions along which we wish to avoid taking positions in our trading portfolio. A simple way to achieve this is to include the K_s vectors $V_{As}^{(a)}$ as columns in the linear constraints matrix $Q_{A\alpha}$ in (4) for good alphas labeled by $i \in J$. I.e., we simply neutralize good alphas w.r.t. $V_{As}^{(a)}$. A simple way to do this is, for each value of $i \in J$ and each value of s, to run a linear regression without the intercept of P_{iAs} over $V_{As}^{(a)}$ and (up to overall normalization factors γ_{is} – see below) to take the residuals of this regression as the new desired holdings, call them \widetilde{P}_{iAs} :

$$\widetilde{P}_{iAs} = \gamma_{is} \left[P_{iAs} - \sum_{a=1}^{K_s} \sum_{B=1}^{M} V_{As}^{(a)} V_{Bs}^{(a)} P_{iBs} \right]$$
(9)

Note that we have

$$\sum_{A=1}^{M} \widetilde{P}_{iAs} \ V_{As}^{(a)} \equiv 0 \tag{10}$$

which is due to the orthogonality of the principal components:

$$\sum_{A=1}^{M} V_{As}^{(a)} V_{As}^{(b)} = \delta_{ab} \tag{11}$$

Here δ_{ab} is the Kronecker delta. The overall normalization factors γ_{is} are fixed via

$$\sum_{A=1}^{M} |\widetilde{P}_{iAs}| = 1 \tag{12}$$

which are the same normalization requirements as (2) for the original positions P_{iAs} .

2.2 A Tweak

Above we compute the matrix X_{ABs} via (8), i.e., using the positions P_{iAs} for dead alphas for one trading day (which is the trading day on which we are going to

¹² Here eRank(Z) of a symmetric semi-positive-definite matrix (which suffices for our purposes here) Z is defined as eRank(Z) = exp(H), where $H = -\sum_{a=1}^{L} p_a \ln(p_a)$ has the meaning of the (Shannon a.k.a. spectral) entropy [Campbell, 1960], [Yang et al, 2005], $p_a = \lambda^{(a)} / \sum_{b=1}^{L} \lambda^{(b)}$, and $\lambda^{(a)}$ are the L positive eigenvalues of Z.

achieve the desired holdings \widetilde{P}_{iAs} for good alphas). If we rebuild the risk model (see below) daily, then this is fine. However, this could also introduce additional twitch (turnover) into our desired trades. To mitigate this, we can average such matrices over some number trading days d (which can – but need not – be the same as d in (5) or (6)). I.e., instead of (8) we take

$$X_{ABs} = \frac{1}{d} \sum_{s'=s}^{s+d-1} \sum_{i \in J'} P_{iAs} P_{iBs}$$
 (13)

The rest goes exactly as above. Let us note that while (8) and (13) at first look "in-sample" as they are based on the data for the date s for which we are computing \widetilde{P}_{iAs} , there is nothing improper about this: P_{iAs} for both dead and good alphas are known in advance of the commencement of trading, that is, they are "previsible".

3 Concluding Remarks

Instead of outright neutralizing the desired holdings w.r.t. $V_{As}^{(a)}$ as above (which amounts to regressing P_{iAs} over $V_{As}^{(a)}$ and taking the residuals, up to overall normalization factors), we can take a different route. We can take good alphas labeled by $i \in J$ and i) either combine them with some weights (see, e.g., [Kakushadze and Yu, 2017a]) and trade the resultant portfolio, or ii) calculate the expected returns for stocks using the expected returns for alphas as in [Kakushadze and Yu, 2017c] and directly trade a portfolio of stocks based on these expected returns (without combining alphas). In the case i) above we get a stock portfolio which can be further optimized, e.g., by maximizing the Sharpe ratio [Sharpe, 1994] or via mean-variance optimization [Markowitz, 1952]. For this (as well as in the case ii) above) we need to construct a risk model covariance matrix – call it Φ_{AB} – for our stock portfolio:¹³

$$\Phi_{AB} = \xi_A^2 \ \delta_{AB} + \sum_{\mu,\nu=1}^F \Omega_{A\mu} \ \phi_{\mu\nu} \ \Omega_{B\nu}$$
 (14)

where ξ_A is the specific (idiosyncratic) risk, $\Omega_{A\mu}$ is the risk factor loadings matrix, and $\phi_{\mu\nu}$ is the factor covariance matrix for F risk factors. So, the idea here is that we can include the K_s vectors $V_{As}^{(a)}$ as columns in $\Omega_{A\mu}$. Then, once we optimize using the so-constructed Φ_{AB} , the resulting stock portfolio is only approximately neutral w.r.t. $V_{As}^{(a)}$. Exact neutrality is attained in the $\xi_A \to 0$ limit [Kakushadze, 2015].

Whether we neutralize or optimize w.r.t. to the directions defined by dead alphas, the upshot is that dead alphas are actually useful. While they no longer make money, if used cerebrally, they help make more money and/or reduce portfolio volatility by eliminating/minimizing taking positions along these direction when trading good alphas. Finally, we present no backtests here as position data is highly proprietary.

¹³ See, e.g., [Grinold and Kahn, 2000] for a general discussion. For an explicit open-source implementation of a general multifactor risk model for equities, see [Kakushadze and Yu, 2016].

A R Source Code

In this appendix we give the R source code¹⁴ for computing the directions associated with dead alphas and factoring them out of good (tradable) alphas. The code below is essentially self-explanatory and straightforward as it simply follows the algorithms and formulas in Section 2. It consists of a single function dead.alphas(hld.good, hld.dead, d, do.trunc = T). Here hld.dead is a 3-dimensional $N_{dead} \times M \times d$ array P_{iAs} of stock positions for dead alphas labeled by $i \in J'$ (see Section 2); d is the number of trading days over which the matrices are averaged in (13); hld.good is a 2-dimensional $N_{good} \times M$ matrix $P_{iA}^* = P_{iAs}$ of stock positions for good alphas labeled by $i \in J$ for the most recent ("today's") date s = 1 for which we wish to compute the desired holdings \widetilde{P}_{iAs} (see Section 2); do.trunc controls whether the eRank is truncated or rounded. Internally the function dead.alphas() calls the function calc.erank(), which is given in Appendix A of [Kakushadze and Yu, 2017c]. The output is an $(N_{good} + K) \times M$ matrix, whose first N_{good} rows are the matrix \widetilde{P}_{iAs} (for s=1), and the remaining rows are the K principal components $V_A^{(a)}$, $a=1,\ldots,K$, where K is computed using eRank (see Section 2). Here we assume that $M \ll N_{dead}$. If the situation is reversed, i.e, we have not too many dead alphas $(N_{dead} \ll M)$, then it is more efficient to replace the built-in R function eigen() (which is based on power iterations [Mises and Pollaczek-Geiringer, 1929) by the function qrm.calc.eigen.eff() given in Appendix C of [Kakushadze and Yu, 2017b, which provides a much more efficient method than power iterations.

```
dead.alphas <- function (hld.good, hld.dead, d, do.trunc = T)
{
    x <- 0
    for(s in 1:d)
        x <- x + t(hld.dead[, , s]) %*% hld.dead[, , s]

    x <- eigen(x)
    x.val <- x$values
    x.vec <- x$vectors
    k <- calc.erank(x.val, excl.first = F)
    if(do.trunc)
        k <- trunc(k)
    else
        k <- round(k)

    q <- x.vec[, 1:k]
    y <- t(hld.good)
    y <- residuals(lm(y ~ -1 + q))</pre>
```

¹⁴ The R Project for Statistical Computing, www.r-project.org.

```
y <- t(y) / colSums(abs(y))
y <- rbind(y, t(q))
return(y)
}</pre>
```

B DISCLAIMERS

Wherever the context so requires, the masculine gender includes the feminine and/or neuter, and the singular form includes the plural and *vice versa*. The author of this paper ("Author") and his affiliates including without limitation Quantigic[®] Solutions LLC ("Author's Affiliates" or "his Affiliates") make no implied or express warranties or any other representations whatsoever, including without limitation implied warranties of merchantability and fitness for a particular purpose, in connection with or with regard to the content of this paper including without limitation any code or algorithms contained herein ("Content").

The reader may use the Content solely at his/her/its own risk and the reader shall have no claims whatsoever against the Author or his Affiliates and the Author and his Affiliates shall have no liability whatsoever to the reader or any third party whatsoever for any loss, expense, opportunity cost, damages or any other adverse effects whatsoever relating to or arising from the use of the Content by the reader including without any limitation whatsoever: any direct, indirect, incidental, special, consequential or any other damages incurred by the reader, however caused and under any theory of liability; any loss of profit (whether incurred directly or indirectly), any loss of goodwill or reputation, any loss of data suffered, cost of procurement of substitute goods or services, or any other tangible or intangible loss; any reliance placed by the reader on the completeness, accuracy or existence of the Content or any other effect of using the Content; and any and all other adversities or negative effects the reader might encounter in using the Content irrespective of whether the Author or his Affiliates is or are or should have been aware of such adversities or negative effects.

The R code included in Appendix A hereof is part of the copyrighted R code of Quantigic[®] Solutions LLC and is provided herein with the express permission of Quantigic[®] Solutions LLC. The copyright owner retains all rights, title and interest in and to its copyrighted source code included in Appendix A hereof and any and all copyrights therefor.

References

Campbell, L.L. (1960) Minimum coefficient rate for stationary random processes. *Information and Control* 3(4): 360-371.

Grinold, R.C. and Kahn, R.N. (2000) Active Portfolio Management. New York, NY: McGraw-Hill.

Kakushadze, Z. (2015) Mean-Reversion and Optimization. *Journal of Asset Management* 16(1): 14-40.

Available online: http://ssrn.com/abstract=2478345.

Kakushadze, Z. (2016) 101 Formulaic Alphas. Wilmott Magazine 2016(84): 72-80. Available online: http://ssrn.com/abstract=2701346.

Kakushadze, Z. and Tulchinsky, I. (2016) Performance v. Turnover: A Story by 4,000 Alphas. *The Journal of Investment Strategies* 5(2): 75-89.

Available online: http://ssrn.com/abstract=2657603.

Kakushadze, Z. and Yu, W. (2016) Multifactor Risk Models and Heterotic CAPM. The Journal of Investment Strategies 5(4): 1-49.

Available online: http://ssrn.com/abstract=2722093.

Kakushadze, Z. and Yu, W. (2017a) How to Combine a Billion Alphas. *Journal of Asset Management* 18(1): 64-80.

Available online: http://ssrn.com/abstract=2739219.

Kakushadze, Z. and Yu, W. (2017b) Statistical Risk Models. *The Journal of Investment Strategies* 6(2): 1-40.

Available online: http://ssrn.com/abstract=2732453.

Kakushadze, Z. and Yu, W. (2017c) Decoding Stock Market with Quant Alphas. *Working Paper*. Available online: https://ssrn.com/abstract=2965224.

Markowitz, H. (1952) Portfolio selection. The Journal of Finance 7(1): 77-91.

Mises, R.V. and Pollaczek-Geiringer. H. (1929) Praktische Verfahren der Gleichungsauflösung. ZAMM – Zeitschrift für Angewandte Mathematik und Mechanik 9(2): 152-164.

Roy, O. and Vetterli, M. (2007) The effective rank: A measure of effective dimensionality. In: *Proceedings – EUSIPCO 2007, 15th European Signal Processing Conference*. Poznań, Poland (September 3-7), pp. 606-610.

Sharpe, W.F. (1994) The Sharpe Ratio. The Journal of Portfolio Management 21(1): 49-58.

Tulchinsky, I. et al. (2015) Finding Alphas: A Quantitative Approach to Building Trading Strategies. New York, NY: Wiley.

Yang, W., Gibson, J.D. and He, T. (2005) Coefficient rate and lossy source coding. *IEEE Transactions on Information Theory* 51(1): 381-386.