人工知能学会研究会資料
SIG-FIN-016-09

なぜロバストなアルファ（ベータ）推定値が必要なのか
ープレミアムVS アノマリーを統計的手法で超えられる？ー

Why We Need Robust Alphas (and Betas)

- Is it possible to beyond “premium vs. anomaly” with the statistical method?

西山　昇

Noboru Nishiyama

Dragons’ Desk Ltd. / 千葉商科大学会計大学院

Abstract: OLS (Ordinary Least Square) estimation of sample means and betas can lead to biased estimates of alpha in the presence of certain patterns of heteroscedasticity. There has been discussing on something to do with background of modern portfolio theory for forecasting risk premium and dealing with anomaly. We demonstrate that those patterns occur in practice, and that a robust estimation process eliminates the bias with some algorithms.

Key words: Alpha, Algorithms, Arbitrage Pricing Theory, Backtesting, Linear Factor Models

1. はじめに

最近、資産運用に人工知能を活用していることを公表する運用者が増加している。ヘッジファンド等は運用の詳細を公開していないため、限られた情報から活用方法を推し量るしかない。少なくとも確実なのは良好なパフォーマンスをあげる（アルファをとる）ためのアルゴリズム構築が中心であることは理解できる。[2]

個社ごとに方法は異なるものの、銘柄選択を通じてプラスアルファが獲得できるポートフォリオ構築のための最適化アルゴリズム開発と先物等のデリバティブを組み合わせたトレーディングの売買タイミングを判断するアルゴリズム開発が中心であろう。

銘柄選択によるプラスアルファの推定モデルを構築するには、ロバストなアルファ値が推定できるモデルと品質の高い原データのインプットが求められる。

CAPM のマーケットモデル（(1）式）を前提とすると株式ポートフォリオの期待リターン($E(r_i)$)の源泉はリスクフリーリターン(r_f)とリスクプレミアム($\beta(\cdot)$)である。

$$r_i = \alpha_i + \beta_i r_M + \epsilon_i \tag{1}$$

$$E(r_i) = r_f + \beta_i \times \left(E(r_M) - r_f\right) \tag{2}$$

リスクフリーリターンは、安全資産（国債、預貯金等）から得られるリターンであり、βがゼロであれば期待リターンとリスクフリーリターンは等しくなる。

裁定価格理論（APT（Arbitrage Pricing Theory）での古典的なノー・アービトラージ仮説は、個別証券のリターンが共通リスクファクターへのエクスポージャの線形関数であるとしている。

ここでのリスクプレミアムとは、株式に投資することによって得られる追加的なリターンである。またアクティブリターンとは、優れた銘柄選択能力により得られる追加的なリターンをさす。

そこではシングルファクター、マルチファクターの差はあるもののリスクプレミアムの係数であるベータの推定方法の精度が重要になる。

これまでリスクとリターンの関係から大きなリスクをとることが大きなリターンを得ることにつながるとされてきた。ところが近年ボラティリティパズルと呼ばれる理論とは逆の現象が指摘されるようになっている[1]。

伝統的な投資理論では、ハイリスク・ハイリターン、ローリスク・ローリターン、なのに対して、観

察される現象は、ハイリスク・ローリターン、ローリスク・ハイリターンとなる場合である。
この要因としてリスクプレミアムが十分に特定されていない説明変数の数を含めた変数選択の問題なのか、アノマリーと呼ばれる一時的な現象であるのかで見解が分かれてきた。
本稿では、Stroyny, Wilding,のリサーチ[3]をもとにリスクプレミアムｖｓアノマリーの議論とは別の視点となる統計的手法によるα、βの推定精度向上について議論する。さらに推定バイアスの存在を確認し、アルファ獲得の可能性を向上させる方法について検討する。

2. 線形ファクターモデル

2.1　アルファとは何か
インデックを対ベンチマークとしてプラスアルファを狙う場合（相対リターン）と投資開始時点よりも資産の評価額が増加する（期初の水準をベンチマークとしてアルファを狙う）場合（絶対リターン）に大きくわけられる。
アウトオブサンプルのデータから将来の期待リターンを予測できれば、その銘柄に対して計算された評価値の高い順番にポートフォリオを組むことでアルファが獲得できるはずである。
筆者はかつてアルファモデルの開発に取り組んだ経験がある。一番のキーとなるのはアルファ値の推定精度である。そのため過去の局面ごとにアルファ値を使ったシミュレーションを行い、上昇相場でも下降相場でもアルファ（相対リターン）を獲得できるモデル開発を目指していた。
パフォーマンスの良し悪しは、アルファ値の予測力に加えて最適化の条件設定にも関連する。アルファ値（期待リターン）を最大化しつつ、ベンチマークから離れないパフォーマンスを実現する最適化は長年の間テーマとなっていた。

2.2　同時推定バイアス（CEB）
APT では個別銘柄の期待リターンは線形ファクタープロセスによって生成されるとしている。
線形ファクターモデルの開発過程では、同じデータセットからアルファ（パフォーマンス評価値）とベータ（リスク指標）を推定するときに発生する同時推定バイアス（CEB）の課題に直面することになる。
この課題はそれに取り組むことによって同時推定バイアス（CEB）を避けるための推定手法の開発につながったとのプラスの側面がある。
ただ同時推定バイアス（CEB）についての基本的な原因については十分に周知されているとは言い難い。
APT では、r は n 銘柄の株式リターンであり、次の線形ファクタープロセスから生成される。

$$r = \mu + Bz + e \qquad (1)$$

ここで B は（n×k）のファクターβの行列、μ は個別銘柄 n 個の期待リターン、z は k 個のシステマティックファクターとなり，e は n 個の残差（銘柄固有）リターンとなる。
定義より E[z] = 0 と E[e] = 0 である。株式固有リターンはお互い無相関であり(E[e,e’] = 0)、システマティックファクターとも無相関 (E[e,z’] = 0)である。
期待リターンとシステマティックファクターリターンとの関係は次のように記述できる。

$$E[r] = \mu = r_0 + B\pi + \alpha \qquad (2)$$

ここで π は k 個のシステマティックファクターのリスクプレミアムとなり、r_0 はリスクフリーリターンである。APT における無裁定の前提は期待リターンとリスクエクスポージャの関係が α ＝ 0 のとき成り立つ。
最近は安定してポートフォリオアルファを獲得することがますますむつかしくなっているマーケット環境から、ここ数年市場ベータ以外のリスクファクターにエクスポージャをもつインデックスにベットして運用するスマートベータとよばれる運用手法に注目が集まる傾向がある。

2.3　最小二乗回帰（OLS）
最小二乗回帰（OLS）は、残差分散が経時的に同一であることを仮定している。しかし実際には Connor, Korajczyk, and Linton (CKL)が議論しているが、残差リターンは時間によって変化する。
そのため彼らは時系列不均一分散を近似したダイナミックなファクターモデル（単共通コンポーネントモデル）を開発している。
ただ、このモデルの仮定はすべての個別銘柄の残差分散が時系列的に同一の動きをすることにある。そのため実際の個別銘柄の残差分散が時系列に不均一となる動きを前提としていない。
個別銘柄の残差分散の変動は、個別銘柄ごと、時間的にも異なる点から、実務的に適用する場合には注意が必要となる。
不均一分散の誤差項のあるパターンから得られる最小二乗回帰（OLS）による標本平均とベータがバ

イアスのあるリスクプレミアム推定値となる例を示す。

これはいわゆる同時推定バイアス（CEB））と呼ばれる。過去の個別銘柄のリターンの不均一分散のパターンを確認すると同時にそのバイアスを消去するロバストな推定プロセスを検討する。

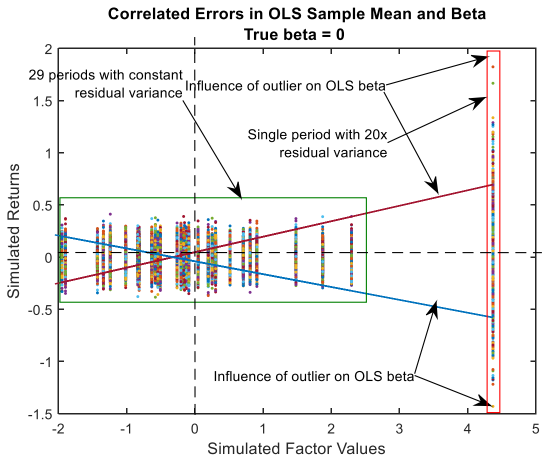
3. 推定バイアス

3.1 外れ値の効果

Stroyny は、最小二乗（OLS）推定における不均一分散をもつ残差、誤差の特定のパターンが標本平均とベータに相関を発生させることを示した。

ベータと標本平均は誤ったリスクリターントレードオフを生成し、それによりアルファの偏った推定値を算出することになる。

実際、線形ファクターモデルにおける最小二乗推定（OLS）の残差が均一分散との仮定と一致しない事例は多い。



（図 1） β を推定する際の外れ値の効果

グラフ（図 1）の Y は 30 変数の正規乱数を 100 期間生成して、次に 1 個のファクターを自由度 3 の T 分布で 100 期間発生させる。

1 個のファクターの残差分散を通常の残差分散の 20 倍として乱数を生成している。残差分散にレバレッジがかかった場合のシミュレーションである。

そこに引かれた回帰直線は、1 期間の外れ値をプラスとしてとる場合には右肩上がりの回帰直線となり、マイナスの外れ値をとる場合には右肩下がりの回帰直線となる。

よって真の β が 0 との前提があるにもかかわらず極端な外れ値があると、その前提が影響を受けることがわかる。

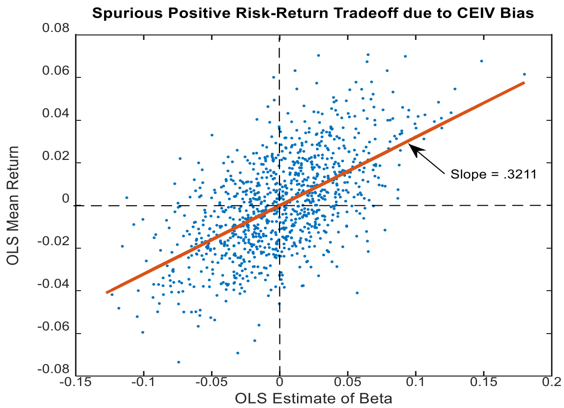
3.2 「相関を持った変数誤差」(CEIV)

単一の過去のデータセットから標本平均とベータの最小二乗（OLS）推定を実行すると、潜在的に各変数の誤差間に相関が生じる Correlated Errors-In-Variables (CEIV)「相関を持った変数誤差」の偏りが生じる。

ベータを推定するための過去データのグループと標本平均を推定する過去データのグループに分割する等の同時推定バイアス（CEB）を避けるための多くの推定方法が考案されてきたが、これらのアプローチはそれぞれのケースで利用可能な情報が半分となり非効率である。

（図 2）では、相関を持った変数誤差の例を示している。横軸に最小二乗回帰による β （リスク）と縦軸に期待リターン（リターン）をとり、データをプロットして回帰直線（OLS）を引くと傾きがプラスとなっている。

リスクプレミアムの推定に本来はリスクとリターンのトレードオフが存在するとの前提関係が、疑似的なプラスのトレードオフ関係が観察される結果となっている。



（図 2）「相関を持った変数誤差」の偏りの例

4. 統計的なソリューション

4.1 収益データ

（図 3）は、時系列プロットである。2 つの分散成分（Scale1, Scale2）の推定結果と米国の四半期の週次の収益リリースに対する反応を示す。

分散成分 1 (Scale1)のピークの値は期間 66 のところであり、2008 年 10 月 10 日 S&P500 指数が週次で最大の下落を記録したときに対応している。グラフをみると 13 週ごとに分散の大きさがピークとなっていることが観察される。業績発表のグラフと重ねてみると分散が上昇するタイミングは収益発表のピ

ークと一致している。

収益発表とは、銘柄固有の情報であり、発表されることにより残差分散が影響を受けることは自然である。

固有ファクターの最大ファクターが 66 週目になっていることから、単一の分散スケールが 3 つの最大ファクターレバレッジにそっているのは先のシミュレーションと平仄があう。

これらのことからリスクプレミアムの OLS 推定による時系列の推定値にはシミュレーションで示した同時推定バイアス（CEB）が存在する可能性が高いと判断できる。

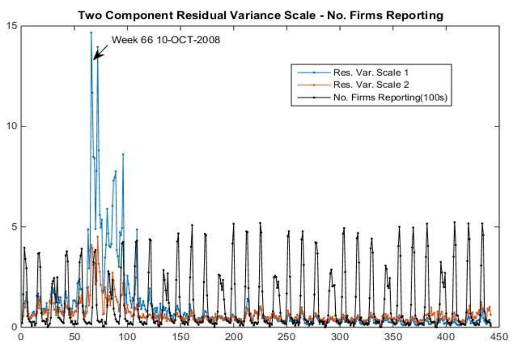
4.2　統計的なソリューション

それでは、上記のバイアスを回避する方法はあるのだろうか。その一つの案としては、OLS（最小二乗回帰分析）の代わりに WLS（重み付き最小二乗回帰分析）を利用する方法である。

（図 3）の Scale1 は Connor, Korajczyk, and Linton (CKL) の単成分分散モデルに週次のデータを適用して近似した結果であり、Scale2 は線形ファクターモデルとして多成分分散モデルを適用した結果である。

多成分分散モデルは、3 個の固有ファクターに 2 成分正規混合残差モデルを組み合わせ最尤法により推定している。

グラフをみると四半期の収益発表が残差分散に影響を与え、単成分分散モデルでは、そのショックを緩和できている。ただ 2008 年 10 月 10 日 S&P500 の下落時には最大の分散の変動を受けている。つまり単一モデルでは、避けることができない残差分散ショックがあることがわかる。



（図 3）2 つの分散成分（Scale1, Scale2）と
米国四半期の週次の収益リリースに対する反応

一方で AL Stroyny が考案した多成分分散モデルでは、2008 年 10 月 10 日のショックを吸収できていることが観察できる。その時点以外の 4 半期の収益発表のショックの影響も緩和できている。

５．おわりに

本稿では、残差分散の変動が β に影響して、α の推定値の精度を低下させている可能性をシミュレーションにより例示した。

そしてそれを解決するためにひとつの統計手法として多成分分散モデルの適用を行っている。

四半期の収益発表の季節変動とそれに加えて大きなマーケットからのショックを含めたアノマリーを統計モデルにより捕捉するひとつの方法を示唆している。

謝辞

Alvin L. Stroyny, Ph.D.と Timothy C. Wilding, Ph.D が彼らの研究成果について情報共有してくれたことに感謝したい。また David Androsoni が本稿の作成にあたりサポートしてくれたことに感謝する。

参考文献

[１] 加藤康之，スマートベーター新時代の投資理論－，応用経済時系列研究会第 23 回談話会資料，2016 年 2 月 16 日

[２] http://www.bloomberg.com/news/articles/2015-02-27/bridgewater-is-said-to-start-artificial-intelligence-team

[３] Stroyny., Wilding., Why Robust Alpha (and Beta)?, discussion paper, unpublished, 〈2016〉