

基礎研 レポート

相関係数を改めて考える

「見方」を変えると答えが変わる

金融研究部 准主任研究員 高岡 和佳子
(03)3512-1851 takaoka@nli-research.co.jp

1—はじめに

資産運用リスクの尺度として分散（収益の散らばり具合）やV a R（バリュー・アット・リスク）といった様々な尺度が提案され、利用されている。どの尺度を利用するかを問わず、複数の資産から構成されるポートフォリオのリスクを定量化するには、個別資産の分散だけでなく、2資産の価格変動（収益率）がどの程度関係しているか（相関係数）を把握することが必要不可欠である。

年次収益率の分散や相関係数が必要な場合、しばしば四半期次や月次といったより短い時間間隔の収益率データを基に年次収益率の分散や相関係数を計算する。これは、年次収益率のように時間間隔の長い収益率では統計的に十分なデータ数を確保できず、信頼に値する推計結果が得られないからである。

日経平均株価が前日比3%弱上昇したある日のニュースでは「大幅反発」と表現されていた。前日比▲7%程度下落した日は、過去歴代何番目の下げ幅かといった記事が出るほど珍しい。しかし、1年を通してみると、日経平均株価が前年比7%の下落（上昇）することは珍しくない。5年程前には1年間で40%以上の下落を経験し、昨年は50%以上の上昇を経験している。このように同じデータ系列であっても、日次収益率の散らばり具合と年次収益率の散らばり具合は明らかに異なる。

このため、十分な数のデータが得られず時間間隔の異なる収益率データで代用する場合は、収益率データから得られた平均、分散、標準偏差などの統計値をそのまま用いることはない。時間間隔の長さの比率を用いて、統計値の特性に応じた加工を施し活用する。しかし、相関係数に関しては、より短い時間間隔の収益率データから得られた相関係数をそのまま利用するのが一般的である。相関係数は2資産の収益率がどの程度関係しているかを表す尺度であり、日次や年次といった時間間隔によらず一定で、統計的な加工は不要と考えられている。

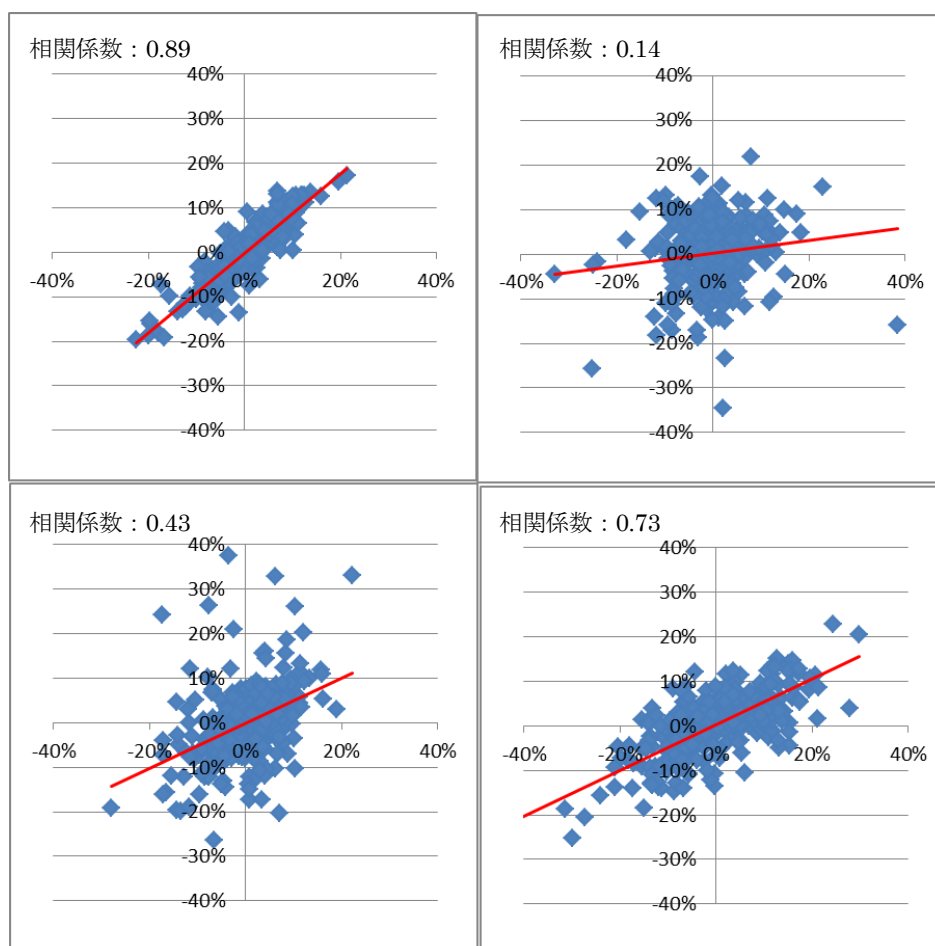
相関係数を推計する際の問題点として経済環境の変化などの影響もあり、時点により結果が異なることがしばしば指摘されるが、時間間隔により結果が異なる可能性についてはあまり指摘されない。

当レポートでは、時間間隔により相関係数が異なる可能性を指摘する。そして、目的とは異なる時間間隔の収益率を用い推計した相関係数を、そのまま利用することの是非を改めて考えることを促す。

2—相関係数とは

相関係数とは、変数間のかかわり合いの程度を表す尺度で、 -1 から 1 の間の値をとる。図表-1はいずれもある2変数をプロットし、赤い線(回帰式)は両者の関係を近似した結果を表している。左上の図のように、一方の変数が大きいほど他の変数も大きい傾向が強ければ 1 に近い値をとり、右上の図のように一方の変数の大きさと、他方の変数の大きさの関係性が小さいと 0 に近い値をとる。なお、相関係数は左上の図が 0.89 、右上の図が 0.14 である。共に赤い線が右肩上がりになっているが、左肩上がりなら相関係数は負の値となる。

図表-1 2変数のプロットと線形近似結果の例



しかし、相関係数が 1 または -1 に近いほど、赤い線の傾きが急になると勘違いしてはいけない。相関係数の符号は右肩上がりか左肩上がりかによって決まるが、傾きの程度と相関係数の大きさは無関係である。左下の図、右下の図の赤い線の傾きはほぼ同程度だが、相関係数は左下の図が 0.43 、右下の図が 0.73 である。相関係数の絶対値を決めるのは、赤い線の傾きではなくプロットの密集度であ

る。基本的には、プロットが基準となる赤い線の近くに密集しているほど相関係数は1または-1に近づく¹。

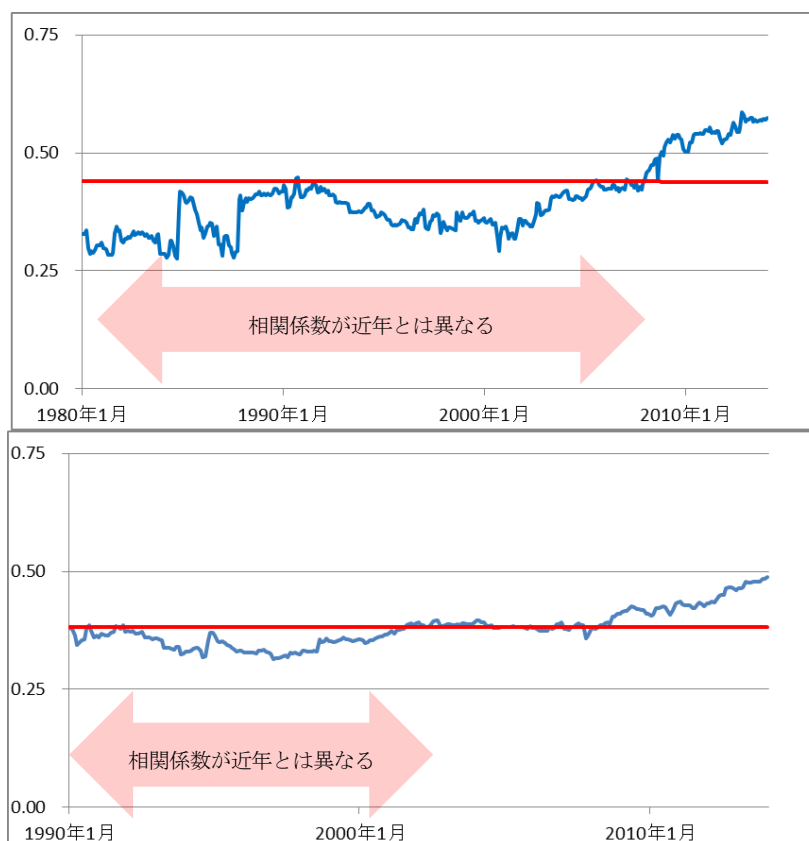
3——相関係数を改めて考える

1 | 相関係数は時点によらず一定か？

本題に入る前に、SP500 収益率と TOPIX 収益率を例に、時点によって相関係数が変わる様子を確認する。図表-2の青線は、推計に用いる期間を1ヶ月ずつずらして推計した相関係数の推移を表している。相関係数に限らず、限られたデータで推計した結果（以下、推定値）は推定に用いるデータ期間を1ヶ月動かすだけで値が変わる。これは、推定値が必ずしも真の相関係数とは限らないことを意味する。しかし、一定の確率で真の相関係数が存在する範囲を求めることが可能だ。赤線は直近における真の相関係数と各時点の真の相関係数が95%の確率で異なると判断可能な分岐点（下回れば異なると判断可能）を表している。表から、直近における真の相関係数と統計的に異なると判断できない期間は短いことが分かる。

図表-2 時点別相関係数の推移（SP500 収益率対 TOPIX 収益率）

（上：直前10年間の月次データより推計、下：直前20年間の月次データより推計）



（注）リーマンショック後の2008年10月を含む期間と、それ以外で相関係数が大きく異なるため、2008年10月の月次収益率データは外して推計している。

¹ 傾きが完全に0の場合は、プロットが基準となる赤線近くに密集していても、相関係数は0となる。

図表－２から、相関係数は時点によって異なることがわかる。SP500 収益率と TOPIX 収益率間の相関係数は総じて上昇傾向にあるが、これは投資のグローバル化による影響と解釈される。

2 | 相関係数は時間間隔によらず一定か？

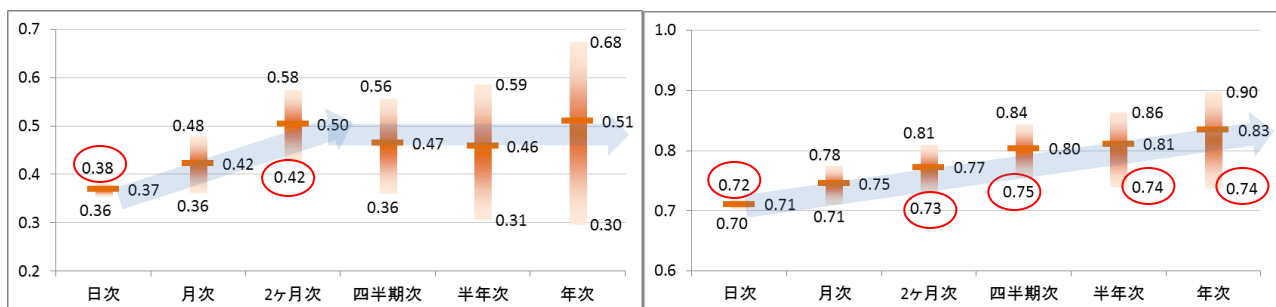
次に、相関係数が時間間隔によらず一定か否かを、SP500 対 TOPIX と TOPIX ガラス・土石指数対 TOPIX 繊維製品指数の 2 組のヒストリカルデータ (1970 年 2 月～2014 年 1 月) を用いて確認する (図表－3)。それぞれ棒の中心付近の印と対応する値は、ヒストリカルデータから推定された相関係数である。棒は、真の相関係数が 90% の確率で存在すると考えられる範囲を表現している。例えば、左図の日次の場合、真の相関係数は 90% の確率で 0.36 から 0.38 の間にあることを示している。裏を返すと、真の相関係数は 0.38 より大きい可能性及び真の相関係数が 0.36 より小さい可能性がそれぞれ 5% あるとも言える。

いずれの組も時間間隔が長くなるほど棒が長いのは、統計的に十分なデータ数を確保できず、信頼に値する推計結果が得られないためである。時間間隔が長くなるにつれ信頼性は低下するとはいえ、いずれの結果も時間間隔が長くなるほど、相関係数が大きくなる傾向に注目してほしい。特に左図の日次データによる結果と 2 ヶ月次データによる結果、及び右図の日次データによる結果と 2 ヶ月次以上の時間間隔のデータによる結果では、真の相関係数が 90% の確率で存在する範囲が重ならない。時間間隔は異なるものの、同じ時点のデータを用いて推計しているため、この相違は 1 節で確認した時点によって相関係数が異なる効果とは別物である。相関係数は日次や年次といった時間間隔によらず一定と考えられるが、この点について再考の余地がありそうだ。

更に、左図は日次データから 2 ヶ月次データまでは相関係数が大きくなっているがそれ以降は収束しほぼ同水準に止まっている。このことから、相関係数の時間間隔が長くなるにつれ相関係数が大きくなるが、一定程度時間間隔が長くなると収束する可能性が伺える。更に右図は収束する傾向が見られないことから、収束スピードはデータによって異なる可能性もうかがえる。

図表－3 時間間隔別相関係数の差 (1970 年 2 月～2014 年 1 月)

(左: SP500 収益率対 TOPIX 収益率、右: TOPIX ガラス・土石指数収益率対 TOPIX 繊維製品指数収益率)



(注) 日次データで SP500 収益率と TOPIX 収益率の相関係数を求める際のみ、同日の収益率ではなく、SP500 収益率と翌営業日の TOPIX 収益率を用いている。同日の収益率を用いると推定値が 0.10、真の相関係数が 90% の確率で存在する範囲の上限、下限がそれぞれ 0.12 と 0.09 となる。

3 | 偶然か？必然か？「見方」を変えると答えが変わる

相関係数は時間間隔によらず一定だという視点で前節の結果を見ると、時間間隔が長くなるほど相関係数が大きくなる傾向も、一定程度時間間隔が長くなると収束する傾向も、また収束スピードがデータによって異なることも、全て、収益率データに限りがあることから生じる、偶然の結果ということになる。真の相関係数が90%の確率で存在する範囲が重ならないことも、一方が真の相関係数が90%の確率で存在する範囲の下限を下回る5%の偶然と、他方が上限を上回ると5%の偶然が同時に起こった稀有の結果ということになる。

しかし、本稿では、稀有の結果として処理しない。相関係数が時間間隔によらず変化する可能性とその理由を探ることで、相関係数は日次や年次といった時間間隔によらず一定と考えることの適切性を再考する。

証券投資理論の多くは市場の効率性を前提とする。市場の効率性は、取引に制限はなく、取引コストや税金もかからないといった取引上の効率性と、証券価格に影響を与える情報が出たらすぐに価格に反映され、また、なんらかの理由で理論価格からの乖離が生じても速やかに解消されるといった価格形成の効率性に大別できる。実際の取引においては取引単位が決まっており、また取引コストや税金もかかるので、明らかに取引上の効率性が満たされているとは言えない。また、取引時間が限られていることに加え、取引の相手(買い手や売り手)がいないために売買が成立しないこともあるのだから、価格形成の効率性も疑わしい。

市場が完全に効率的だという仮定の下では、取引価格と理論価格は一致する。しかし、取引上の非効率性が取引価格と理論価格との乖離を招き、更に価格形成の非効率性により乖離は速やかに解消されない。このように、市場が完全に効率的でなければ、観測される収益率に非効率性がもたらす収益率が含まれる²。

$$\text{(観測される収益率)} = \text{(理論上の収益率)} + \text{(非効率性の収益率)}$$

また、市場が完全に効率的ではないとはいえ乖離はいずれ解消されると考えれば、非効率性がもたらす収益率は過去の非効率性がもたらす収益率を打ち消す効果をもつと考えられる。そこで、非効率性がもたらす収益率と前営業日のその関係は以下の関係式で表現できると仮定する。

² 例えば、理論価格が下表のように推移した場合、売買単位が1円ならば小数点以下で売買することはできない。このため、下表の下段のように推移したと考えるとわかりやすい。

	1 営業日目		2 営業日目		3 営業日目
理論価格	50 円		51.2 円		52.1 円
理論上の収益率		+2.40%		+1.76%	
取引価格	50 円		51 円		52 円
観測される収益率		+2.00%		+1.96%	
非効率性の収益率		▲0.40%		+0.20%	

$$\begin{aligned} \text{(非効率性の収益率)} &= \beta \times \text{(前営業日の非効率性の収益率)} \\ &\quad + \text{(非効率性の収益率のうち当営業日に帰属する部分)} \\ \beta &: \text{解消するスピードを表す係数 } (\beta < 0) \end{aligned}$$

市場が効率的ではないことを踏まえ、上の2つの関係を仮定すると、理論上の収益率間の相関係数が時間間隔によらず一定であったとしても、時間間隔が長くなるほど観測される収益率の相関係数は大きくなる。

なぜならば、前章で説明したように相関係数の絶対値を決めるのは傾きではなくプロットの密集度である。そして、非効率性がもたらす収益率はプロットをばらけさせる効果を持つ。つまり、市場が完全に効率的でないならば観測される収益率間の相関係数（絶対値）は、理論上の収益率間の相関係数（絶対値）より小さくなる。更に、非効率性がもたらす収益率に過去の非効率性がもたらす収益率を打ち消す効果があるのであれば、時間間隔が長いほど非効率性がもたらす収益率の影響は小さくなり、観測される収益率間の相関係数が理論上の収益率間の相関係数に近づくと考えられる。これにより、時間間隔が長くなるほど相関係数が大きくなる現象を説明できる。

また、時間間隔が長くなり非効率性の影響がなくなるにつれ、理論上の相関係数に近づくのだから、一定程度時間間隔が長くなると収束する傾向も説明できる³。更に、解消するスピードを表す係数 β の大きさによって、非効率性の影響がなくなるスピードが異なるため、収束スピードがデータによって異なることも説明ができるかもしれない。このように考えると、前節の結果は偶然ではなく必然ということになる。

4—まとめ

市場が完全に効率的でないならば、時間間隔によって相関係数が異なりうる。つまり、統計的に十分なデータ数を確保できないことを理由に、目的より短い時間間隔の収益率で代用する場合、相関係数の過小評価の可能性に注意する必要がある。幸い、一定程度時間間隔が長くなると相関係数は収束すると考えられる。しかし、収束するスピードは、理論価格からの乖離を解消するスピードに依存する可能性がある。取引が盛んでない資産ほど理論価格からの乖離を解消するスピードが遅いと考えられるので、流動性の乏しい資産間の相関係数を求める際は、特に注意が必要だろう。

当レポートでは、時間間隔によって相関係数が異なる可能性を指摘し、目的とは異なる時間間隔の収益率を用い推計した相関係数を、そのまま利用することの是非を改めて考えることを促すことにとどめる。3章3節のモデルの具体的説明や、複数の時間間隔の収益率を用いて推計した相関係数や分散から長期の収益率間相関係数を見積もる方法の提案は、次のレポートに記載することとする。

³ 実際は、非効率性の影響がなくなることはないが、 β と非効率性の収益率の分散に応じ決まる値に収束する。