人工知能学会研究会資料
SIG-FIN-024

時系列モデルを用いたマルチアセット市場における統計的裁定戦略

Statistical Arbitrage Strategy in Multi-Asset Market using Time Series Analysis

今井　崇公[1*]　中川　慧[1]
Takahiro Imai[1]　Kei Nakagawa[1]

[1] 野村アセットマネジメント株式会社
[1] Nomura Asset Management Co.,Ltd.

Abstract: The statistical arbitrage strategy is one of the most traditional investment strategies. Many theoretical and empirical studies have been conducted for a long time. Almost all of the statistical arbitrage strategies focus on the price difference (spread) between two similar assets in same asset class and exploit the mean reversion of spreads, i.e., pairs trading. In this study, we extend the strategy to multiple assets in the multi-asset market. Concretely, we derive a mean-reverting portfolio with time series model. Finally, we perform an empirical analysis in multi-asset market and show the effectiveness of our strategy.

1　はじめに

資産間の価格差（スプレッド）の平均回帰性に着目した研究は古くから数多く行われており、その性質を利用したペア・トレード戦略は現在でも個人投資家から機関投資家まで幅広い層に利用されている [1]。[2] によると、ペア・トレード戦略を行うにあたってのアプローチは多岐にわたり、銘柄間距離、共和分を代表とする時系列モデル、確率制御を利用するものなどが存在する。一方で、ペア・トレード戦略の多くが類似した銘柄間のスプレッドに着目した戦略であることから、ほとんどの先行研究は単一資産内の銘柄をユニバースとしている。

そこで本研究では、単一資産ではなくマルチアセット市場を対象に時系列モデルを用いて平均回帰性を有するポートフォリオを構築することで、異なる資産クラスの銘柄をロング、ショートするペア・トレード戦略を提案する。マルチアセット市場を対象とすることで、単一資産に留まらないグローバルなアセットクラス間の連動性のずれを裁定機会として得ることが狙いである。具体的には、異なる観点から定義された平均回帰性を定量化する基準を複数用意し、それら複数の基準で見て「良い」ポートフォリオを構築する。その際に、Polynomial Goal Programming(PGP) と呼ばれる多目的最適化問題の手法を用いることで、平均回帰性の定量基準を恣意性なく組み合わせる手法を提案する。そして当該ポートフォリオが平均水準から一定程度乖離したタイミングでポジションをとって、平均回帰する方向にベットする。

本稿の構成は次の通りである。2 章では、関連する時系列モデルを用いた平均回帰ポートフォリオの先行研究について簡単に述べる。3 章では、平均回帰性の「良さ」を示す基準を複数導入すると共に、それらを統合し、最適な保有比率を求める手法である PGP について説明する。そして 4 章において、本研究で対象とするマルチアセット市場におけるペア・トレード戦略を記述し、5 章で実際の金融市場のデータを用いた実証分析によってその有効性を評価する。最後に、結論を述べる。

2　先行研究

平均回帰性の定量基準は様々な形で提案されている。有名な定量基準としては、時系列が分散で測ってどれだけホワイトノイズに近いかを示す Predictability[3, 4]、時系列が相関で測ってどれだけホワイトノイズに近いかを示す Portmanteau Statistics[5]、時系列がある単位時間区間において平均水準に何回回帰したかを示す Crossing Statistics[6] などが存在する。

*連絡先：野村アセットマネジメント株式会社
〒 103-0027　東京都中央区日本橋 1-11-1
E-mail: t-imai@nomura-am.co.jp
本稿の内容は筆者らが所属する組織を代表するものではなく、本稿の全ての誤りは、筆者らの責に属するものである。

これらの指標を用いたペア・トレード戦略として、[7] では米国株式のインプライドボラティリティを、[8] では米国株式を投資対象に Predictability、Portmanteau Statistics、Crossing Statistics をそれぞれ単独に用いて平均回帰ポートフォリオを構築している。 これらの研究は単体の平均回帰性の定量基準の有効性を評価できているという点で有用ではあるが、多角的な視点で平均回帰ポートフォリオを構築できてはいない。ここで複数の平均回帰性の基準を組み合わせる方法として、サブ目的関数を結合した多目的最適化問題を解く方法が考えられるが、最適化の結果に大きく関わる結合係数の決定方法について恣意性が残るという課題がある。このような多目的最適化問題のハイパーパラメータ設定を容易にする方法として、PGP と呼ばれる手法が高次モーメントを含めたポートフォリオ最適化問題においてしばしば用いられている [9]。

3 平均回帰ポートフォリオの構築手法

N 個の資産が存在する時、その時点 t における対数価格を $\boldsymbol{y_t} = \{y_{1,t}, ..., y_{N,t}\}$ とする。各資産のウェイトベクトルを $\boldsymbol{w} = \{w_1, ..., w_N\}$ とすると、ポートフォリオは

$$z_t = \boldsymbol{w}^\top \boldsymbol{y_t} \tag{1}$$

と書け、ポートフォリオの対数収益率 r_t は、

$$r_t = z_t - z_{t-1} = \boldsymbol{w}^\top (\boldsymbol{y_t} - \boldsymbol{y_{t-1}}) \tag{2}$$

と表せる。本研究における問題設定は、式 (1) のポートフォリオ z_t が平均回帰するようなウェイト $\boldsymbol{w}$ を求めることである。すなわち、複数資産間においてできるだけ平均回帰するようなウェイトを算出することが目的である。

3.1 平均回帰の基準

ここでは、ポートフォリオ z_t の平均回帰の良さを定量的に示す基準を複数導入する。本研究では、(1)Predictability、(2)Portmanteau Statistics、(3)Crossing Statistics の 3 つの観点で平均回帰性を定量化する。はじめに、系列 $\boldsymbol{y_t}$ に対する次数 (ラグ)$i \in \mathbb{N}$ の自己共分散行列を次の通り定義する。

$$\begin{aligned} \boldsymbol{M_i} &:= \mathrm{Cov}\left(\boldsymbol{y_t}, \boldsymbol{y_{t+i}}\right) \\ &= \mathbb{E}\left[\left(\boldsymbol{y_t} - \mathbb{E}[\boldsymbol{y_t}]\right)\left(\boldsymbol{y_{t+i}} - \mathbb{E}[\boldsymbol{y_{t+i}}]\right)^\top\right] \end{aligned} \tag{3}$$

式 (3) から明らかであるが、$i = 0$ の $\boldsymbol{M_0}$ は通常の共分散行列を表す。

3.1.1 Predictability

この基準は時系列が分散で測ってどれだけホワイトノイズに近いかという観点から平均回帰性を定量的に表す。定常な以下の時系列を考える。

$$y_t = \hat{y}_{t-1} + \varepsilon_t \tag{4}$$

ここで、$\hat{y}_{t-1}$ は時刻 $t-1$ までの情報に基づく予測値を表し、最も単純な例として $\hat{y}_{t-1}$ を αy_{t-1} とする AR(1) モデルがある。ε_t は $\hat{y}_{t-1}$ とは独立の分散 σ_ε^2 のホワイトノイズを表す。式 (4) の分散をとると、$\sigma_y^2 = \sigma_{\hat{y}}^2 + \sigma_\varepsilon^2$ となる。ここで、σ_y^2 は y_t の分散、$\sigma_{\hat{y}}^2$ は $\hat{y}_{t-1}$ の分散をそれぞれ表す。

Predictability は、

$$predictability = \frac{\sigma_{\hat{y}}^2}{\sigma_y^2} \tag{5}$$

と定義される。

Predictability の定義から、この値が小さいと、ホワイトノイズに近いことを意味し、平均回帰しやすい。一方で値が大きいと、$t-1$ までの情報で予測しやすいことを意味する。

$\hat{y}_{t-1}$ を次の VAR(1) モデルでモデル化できるとする。なお、[10] から一般に VARMA(p, q) モデルは VAR(1) モデルで表現可能であるため、以下の議論は VARMA(p, q) モデルに対しても同様に可能である。

$$\boldsymbol{y_t} = \boldsymbol{A}\boldsymbol{y_{t-1}} + \boldsymbol{e_t} \tag{6}$$

ここで $\boldsymbol{e_t}$ はホワイトノイズである。

式 (6) の VAR(1) モデルに $\boldsymbol{w}^\top$ をかけると、$\boldsymbol{w}^\top\boldsymbol{y_t} = \boldsymbol{w}^\top\boldsymbol{A}\boldsymbol{y_{t-1}} + \boldsymbol{w}^\top\boldsymbol{e_t}$ となる。この分散を取ると、左辺は $\boldsymbol{w}^\top\boldsymbol{M_0}\boldsymbol{w}$ となり、右辺第一項 $\boldsymbol{w}^\top\boldsymbol{A}\boldsymbol{y_{t-1}}$ は $\boldsymbol{w}^\top\boldsymbol{A}\boldsymbol{M_0}\boldsymbol{A}^\top\boldsymbol{w}$ となる。さらに VAR(1) モデルの性質から $\boldsymbol{A} = \boldsymbol{M_1}^\top\boldsymbol{M_0}^{-1}$ であるので、これを代入して結局、右辺第一項は $\boldsymbol{w}^\top\boldsymbol{A}\boldsymbol{M_0}\boldsymbol{A}^\top\boldsymbol{w} = \boldsymbol{w}^\top\boldsymbol{M_1}^\top\boldsymbol{M_0}^{-1}\boldsymbol{M_0}\left(\boldsymbol{M_1}^\top\boldsymbol{M_0}^{-1}\right)^\top\boldsymbol{w} = \boldsymbol{w}^\top\boldsymbol{M_1}^\top\boldsymbol{M_0}^{-1}\boldsymbol{M_1}\boldsymbol{w}$ となる。したがって、VAR(1) モデルの Predictability は、

$$predictability(\boldsymbol{w}) = \frac{\boldsymbol{w}^\top\boldsymbol{M_1}^\top\boldsymbol{M_0}^{-1}\boldsymbol{M_1}\boldsymbol{w}}{\boldsymbol{w}^\top\boldsymbol{M_0}\boldsymbol{w}} \tag{7}$$

となる。

3.1.2 Portmanteau Statistics

この基準は時系列が自己相関で測ってどれだけホワイトノイズに近いかという観点から平均回帰性を定量的に表す。

定常なラグ p を持つ y_t に対して、

$$y_t = \hat{y}_{t-1} + ...\hat{y}_{t-p} + \varepsilon_t \tag{8}$$

Portmanteau Statistics は

$$portmanteau = \sum_{i=1}^{p} \rho_i^2 \tag{9}$$

とかける。ここで ρ_i は、ラグ i の自己相関係数であり、$\mathbb{E}[z_t z_{t+i}]/\mathbb{E}[z_t^2]$ と定義される。式 (10) の定義から $portmanteau \geq 0$ であることがわかり、またホワイトノイズはどのラグ p ととっても自己相関を持たないので、$portmanteau$ が 0 に近ければ、自己相関で測ってホワイトノイズに近いと言える。したがってこれを最小化するように $\boldsymbol{w}$ をとると平均回帰しやすい。

$z_t = \boldsymbol{w}^\top \boldsymbol{y_t}$ の時、自己相関係数は $\rho_i = \mathbb{E}\left[z_t z_{t+i}\right]/\mathbb{E}\left[z_t^2\right]$ より、平均回帰ポートフォリオの portmanteau は次のように書ける。

$$portmanteau(\boldsymbol{w}) = \sum_{i=1}^{p} \left(\frac{\boldsymbol{w}^\top \boldsymbol{M_i} \boldsymbol{w}}{\boldsymbol{w}^\top \boldsymbol{M_0} \boldsymbol{w}} \right)^2 \tag{10}$$

3.1.3 Crossing Statistics

この基準は時系列がある時間区間 T において平均水準に何回回帰したかをカウントし、その数で平均回帰性を定量的に表す。

Crossing statistics は、定常時系列 y_t に対して、以下の通り定義される。

$$crossing = \frac{1}{T-1} \sum_{t=2}^{T} \mathbf{1}_E\left(y_t\right) \tag{11}$$

ここで、$\mathbf{1}_E\left(y_t\right)$ は、事象 $E = \{y_t y_{t-1} \leq 0\}$ を満たす時に 1、満たさない場合には 0 を返す指示関数である。

定常な正規ノイズを持つ時系列に対しては、

$$crossing = \frac{1}{\pi} \arccos\left(\rho_1\right) \tag{12}$$

で与えられることが知られている。この式から、定常な正規ノイズを持つ時系列では、平均水準に多く回帰させるためには、ρ_1 を小さくすることが必要である。

以上から $z_t = \boldsymbol{w}^\top \boldsymbol{y_t}$ の時、平均回帰ポートフォリオの Crossing statistics は 1 次の自己相関の定義から次のように書ける。

$$crossing(\boldsymbol{w}) = \frac{\boldsymbol{w}^\top \boldsymbol{M_1} \boldsymbol{w}}{\boldsymbol{w}^\top \boldsymbol{M_0} \boldsymbol{w}} \tag{13}$$

3.2 Polynomial Goal Programming による定式化

ここでは、前節で導入した 3 つの平均回帰性の定量基準を PGP により統合し、最適な保有比率を決定する。この手法は、高次モーメントを含めたポートフォリオの多目的最適化問題において導入された手法で、複数のサブ目的関数を規格化する点と、それらのサブ目的関数に投資家の選好を反映できる点が利点である。

PGP では、まずサブ目的関数である式 (7)、式 (10)、式 (13) の最小化問題を独立に解き、各々の最小値 $pred^*$、$port^*$、$cross^*$ を求める。次に、それらを用いて

$$\begin{aligned} \underset{\boldsymbol{w}}{\arg\min} \quad & \left|\frac{d_1}{pred^*}\right|^{\lambda_1} + \left|\frac{d_2}{port^*}\right|^{\lambda_2} + \left|\frac{d_3}{cross^*}\right|^{\lambda_3} \\ \text{where} \quad & d_1 = predictability(\boldsymbol{w}) - pred^* \\ & d_2 = portmanteau(\boldsymbol{w}) - port^* \\ & d_3 = crossing(\boldsymbol{w}) - cross^* \end{aligned} \tag{14}$$

を解くことで、最適な $\boldsymbol{w}$ を求める。ここで λ_1、λ_2、λ_3 はそれぞれ Predictability、Portmanteau Statistics、Crossing Statistics の選好度を表す変数である。

3 つの平均回帰性に関する定量基準と PGP による統合について、概念図を図 1 に示す。

4 提案手法

本研究では、前章で導入した定量基準を基にポートフォリオを構築し、移動平均からの乖離が平均回帰する性質を利用したペア・トレード戦略を提案する。

Step 1：

マルチアセットを対象に、あらかじめ設定した最大次数以下で VAR(p) モデルの AIC が最小となる次数 p を選択する。この次数を用いて、各平均回帰性に関する定量基準の値を算出する。

Step 2：

(1)Predictability、(2)Portmanteau Statistics、(3)Crossing Statistics、(4)PGP を用いて (1)〜(3) を統合した多目的関数をそれぞれ最小化することで、平均回帰性を有するポートフォリオのウェイトを算出する。

Step 3：

Step 2 で算出したポートフォリオの過去リターンの移動平均からのスプレッドを算出する。スプレッドが平均 ±1 標準偏差以上乖離した時をシグナルとしてポジションを取得する。またロスカットとして、ポジション取得後に平均 ±2 標準偏差以上乖離した場合は、ポジションを解消する。

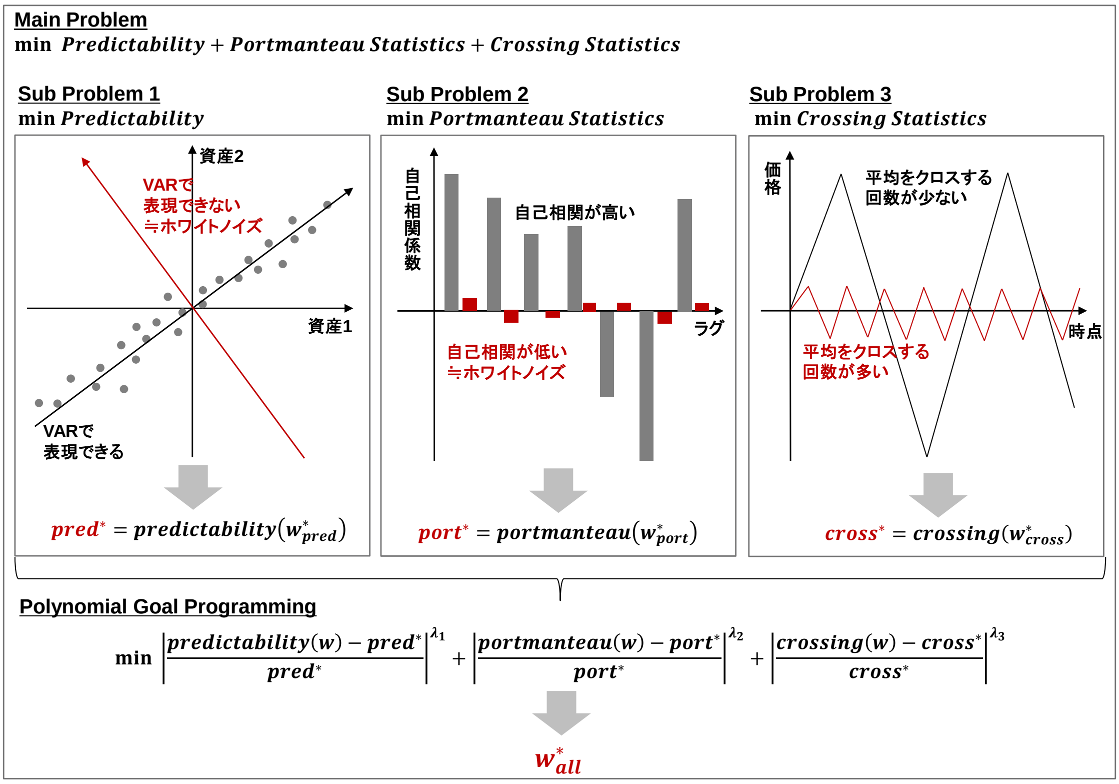

図 1: 3 つの平均回帰性に関する定量基準と PGP による統合

5　実証分析

5.1　データセット

ここではグローバルの先物を対象に分析を行った。投資対象のユニバースを表 1 にまとめた。データは Bloomberg から取得した。検証期間は、2007 年 3 月 23 日から 2019 年 8 月 30 日までとし、すべて金曜引けの週次リターンを用いた。

5.2　分析方法

四半期毎に、最大次数を 5、データ期間を過去 52 週として VAR(p) モデルの次数を決定し、最適なポートフォリオのウェイトを算出した。ここでポートフォリオのレバレッジは、過去 52 週のリターンデータを用いて推定した共分散行列を基に、推定リスクが年率 5%となるように決定した。なお Predictability のみ、PGP で統合することを踏まえ VAR(1) モデルを用いた。仮にポジションを取得していたとしても四半期毎に解消することとした。また、ロスカットを行った場合、当該四半期はポジションを取得しないこととした。PGP における Predictability、Portmanteau Statistics、Crossing Statistics の選好度はすべて 1 とした。ポートフォリオの過去リターンの移動平均は過去 13 週、スプレッドの平均および標準偏差は過去 52 週の値を用いて算出した。以後、ポジション取得から解消までの期間を 1 戦略と呼称する。

5.3　分析結果

各平均回帰性の定量基準を用いて、ポートフォリオを構築した際のパフォーマンス統計量のサマリーを表 2 に、累積リターンの比較を図 2 に示す。

表 2 の各パフォーマンス統計量の定義について簡単に説明する。投資期間のリターン、リスク、リターン/リスクは、それぞれポジションを取得している時のみのポートフォリオのリターン系列を用いて算出しており、リターンおよびリターン/リスクは値が大きいほど、リスクは値が小さいほどパフォーマンスが良いことを示す。戦略の勝率は、ポートフォリオのリターンが正であった戦略数を総戦略数で除した値とし、値が大きいほどパフォーマンスが良いことを示す。戦略の最大ドローダウンは、全戦略のうち最も低いリターンの値を表し、絶対値が小さいほどパフォーマンスが良いことを示す。全期間に占める投資期間の割合は、全期間のうちポジションを取得している期間の割合を表し、1 戦略あたりの平均投資期間は、ポジションを取得してから解消するまでにかかった期間の平均を表す。

表 1: 投資対象資産

	銘柄							
株式先物（計 16）	S&P 500（SP）	NAS DAQ（NQ）	カナダ（PT）	英（CF）	仏（CF）	独（GX）	欧州（VG）	スペイン（IB）
	オランダ（EQ）	ノルウェー（OI）	スイス（SM）	日経（NK）	TOPIX（TP）	香港（HI）	豪州（XP）	シンガポール（QZ）
債券先物（計 13）	米国 10 年（TY）	米国 2 年（TU）	米国 5 年（FV）	米国 20 年（US）	豪州 3 年（YM）	豪州 10 年（XM）	カナダ 10 年（CN）	独 2 年（DU）
	独 5 年（OE）	独 10 年（RX）	独 30 年（UB）	英 10 年（G）	日本 10 年（JB）			

表 2: パフォーマンス統計量

	PGP	Predictability	Portmanteau	Crossing
投資期間のリターン（%、年率換算）	**11.9**	9.9	3.1	2.8
投資期間のリスク（%、年率換算）	6.9	8.9	**4.5**	5.6
投資期間のリターン/リスク	**1.73**	1.11	0.70	0.50
戦略の勝率（%）	**59.8**	57.3	50.6	49.4
戦略の最大ドローダウン（%）	-2.5	-4.4	**-1.7**	-2.7
全期間に占める投資期間の割合 （%）	28.0	24.5	30.0	23.8
1 戦略あたりの平均投資期間（週）	1.9	1.8	2.4	2.0

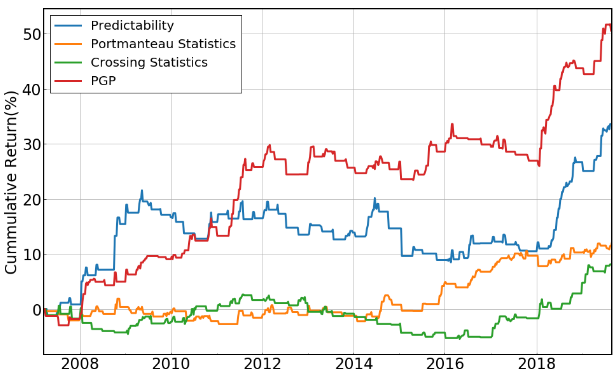

図 2: 累積リターンの比較

表 2 の投資期間のリターン/リスクおよび戦略の勝率では PGP が 3 つの単体手法を優越する結果を示しており、複数の平均回帰性の基準を組み合わせることの有効性を示唆している。図 2 によれば、各単体手法が有効な期間は概ね異なるが、PGP を用いることでこれらをうまく統合できていると言える。

6 まとめ

本研究では、マルチアセット市場を対象に時系列モデルを用いて平均回帰性を有するポートフォリオを構築することで、異なる資産クラスの銘柄をロング、ショートするペア・トレード戦略を提案した。具体的な平均回帰性の定量基準としては、分散の観点で測った Predictability、相関の観点で測った Portmanteau Statistics、平均水準への回帰回数で測った Crossing Statistics、およびこれらを PGP と呼ばれる多目的最適化問題の手法により統合した基準を用いた。実証分析では、PGP を用いた手法において最良のリターン/リスクおよび勝率の結果が得られ、マルチアセット市場のペア・トレード戦略において複数の平均回帰性の定量基準を組み合わせる戦略の有効性が示唆された。

参考文献

[1] Ganapathy Vidyamurthy. *Pairs Trading: quantitative methods and analysis*, volume 217. John Wiley & Sons, 2004.

[2] Christopher Krauss. Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, 31(2):513–545, 2017.

[3] George EP Box and George C Tiao. A canonical analysis of multiple time series. *Biometrika*, 64(2):355–365, 1977.

[4] Ronald Bewley, David Orden, Minxian Yang, and Lance A Fisher. Comparison of box—tiao and johansen canonical estimators of cointegrating vectors in vec (1) models. *Journal of Econometrics*, 64(1-2):3–27, 1994.

[5] Greta M Ljung and George EP Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

[6] N Donald Ylvisaker. The expected number of zeros of a stationary gaussian process. *The Annals of Mathematical Statistics*, 36(3):1043–1046, 1965.

[7] Marco Cuturi and Alexandre d'Aspremont. Mean reversion with a variance threshold. In *International Conference on Machine Learning*, pages 271–279, 2013.

[8] Ziping Zhao and Daniel P Palomar. Mean-reverting portfolio with budget constraint. *IEEE Transactions on Signal Processing*, 66(9):2342–2357, 2018.

[9] Mehmet Aksaraylı and Osman Pala. A polynomial goal programming model for portfolio optimization based on entropy and higher moments. *Expert Systems with Applications*, 94:185–192, 2018.

[10] Helmut Lütkepohl. *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.